510

THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

JUN 25 1951

Volume 2 *No. 6* *June* 1951

S. N. Afriat: Bounds for the Characteristic Values of Matrix Functions 81

G. Szekeres: An Asymptotic Formula in the Theory of Partitions 85

Sheila Scott MacIntyre: Overconvergence Properties of some Interpolation Series 109

S. N. Afriat: An Iterative Process for the Numerical Determination of Characteristic Values of certain Matrices 121

W. H. Cockcroft: The Word Problem in a Group Extension 123

F. F. Bonsall: Inequalities with Non-conjugate Parameters 135

E. M. Patterson: On Symmetric Recurrent Tensors of the Second Order 151

W. H. Cockcroft: Note on a Theorem by J. H. C. Whitehead 159

OXFORD
AT THE CLARENDON PRESS

Price 10*s.* 6*d. net*

PRINTED IN GREAT BRITAIN BY CHARLES BATEY AT THE UNIVERSITY PRESS, OXFORD

THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

Edited by T. W. CHAUNDY, U. S. HASLAM-JONES, J. H. C. THOMPSON

THE QUARTERLY JOURNAL OF MATHEMATICS (OXFORD SECOND SERIES) is published at 10s. 6d. net for a single number with an annual subscription (for four numbers) of 37s. 6d. post free.

Papers, of a length normally not exceeding 20 printed pages of the Journal, are invited on subjects of Pure and Applied Mathematics, and should be addressed 'The Editors, Quarterly Journal of Mathematics, Clarendon Press, Oxford'. They should be in a condition fit for sending to a referee and in form and notation convenient for printing. The Editors as a rule will not wish to accept material that they cannot see their way to publish within a twelvemonth.

While every care is taken of manuscripts submitted for publication, the Publisher and the Editors cannot hold themselves responsible for any loss or damage. Authors are advised to retain a copy of anything they may send for publication. Authors of papers printed in the Quarterly Journal will be entitled to 50 free offprints.

Correspondence on the *subject-matter* of the Quarterly Journal should be addressed, as above, to 'The Editors', at the Clarendon Press. All other correspondence should be addressed to the Publisher

GEOFFREY CUMBERLEGE

OXFORD UNIVERSITY PRESS

AMEN HOUSE, LONDON, E.C. 4

BOUNDS FOR THE CHARACTERISTIC VALUES OF MATRIX FUNCTIONS

By S. N. AFRIAT (*Oxford*)

[Received 13 January 1950]

1. Introduction

For any square complex matrix of finite or infinite order, a pair of non-negative real numbers, called the *upper* and *lower moduli* of the matrix, is introduced. These are upper and lower bounds for the moduli of the characteristic values of the matrix. They generalize the concept of the modulus of a complex number, and satisfy certain inequalities for sums and products.

Conditions for the convergence of infinite matrix sums and products are deduced, and it is shown that an upper bound for the moduli of the characteristic values of a matrix function, formed from sums and products in a set of matrix arguments, is the corresponding scalar function of the upper moduli of the arguments.

2. Definition. *The upper and lower moduli* $|\mathbf{a}|^*$, $|\mathbf{a}|_*$ *of a matrix* $\mathbf{a}$ *are the non-negative real numbers whose squares are the maximum and minimum characteristic values respectively of the non-negative definite Hermitian matrix* $\bar{\mathbf{a}}'\mathbf{a}$.

3. Theorem. *If* α *is a characteristic value of* $\mathbf{a}$, *then*

$$|\mathbf{a}|_* \leqslant |\alpha| \leqslant |\mathbf{a}|^*.$$

Let $\mathbf{x}^*$, $\mathbf{x}_*$ be unit latent vectors of $\bar{\mathbf{a}}'\mathbf{a}$ corresponding to the characteristic values $|\mathbf{a}|^{*2}$, $|\mathbf{a}|_*^2$ respectively. Then the Hermitian form $\bar{\mathbf{x}}'\bar{\mathbf{a}}'\mathbf{a}\mathbf{x}$, subject to the condition $\bar{\mathbf{x}}'\mathbf{x} = 1$, has maximum and minimum values $|\mathbf{a}|^{*2}$, $|\mathbf{a}|_*^2$ attained at $\mathbf{x}^*$, $\mathbf{x}_*$ respectively, as is clear by the method of Lagrangian multipliers. Thus

$$|\mathbf{a}|_*^2 \leqslant \bar{\mathbf{x}}'\bar{\mathbf{a}}'\mathbf{a}\mathbf{x} \leqslant |\mathbf{a}|^{*2}, \quad \text{when } \bar{\mathbf{x}}'\mathbf{x} = 1.$$

Now let $\mathbf{u}$ be a unit latent vector of $\mathbf{a}$ corresponding to a characteristic value α. Then $\mathbf{a}\mathbf{u} = \alpha\mathbf{u}$, and $\bar{\mathbf{u}}'\bar{\mathbf{a}}'\mathbf{a}\mathbf{u} = \bar{\alpha}\alpha\bar{\mathbf{u}}'\mathbf{u} = \bar{\alpha}\alpha$ since $\bar{\mathbf{u}}'\mathbf{u} = 1$; whence, taking $\mathbf{x} = \mathbf{u}$ above,

$$|\mathbf{a}|_*^2 \leqslant \bar{\alpha}\alpha \leqslant |\mathbf{a}|^{*2}.$$

4. Theorem. *For any matrices* $\mathbf{a}$, $\mathbf{b}$ *and a scalar matrix* $\mathbf{k}$ $(= k\mathbf{1}$ *where* k *is any scalar*),

$$|\mathbf{k}|^* = |\mathbf{k}|_* = |k|; \qquad |\mathbf{k}\mathbf{a}|^* = |k||\mathbf{a}|^*, \qquad |\mathbf{k}\mathbf{a}|_* = |k||\mathbf{a}|_*;$$

$$|\mathbf{a}^{-1}|^* = (|\mathbf{a}|_*)^{-1}, \qquad |\mathbf{a}^{-1}|_* = (|\mathbf{a}|^*)^{-1};$$

$$|\mathbf{a}|_*|\mathbf{b}|_* \leqslant |\mathbf{a}\mathbf{b}|_* \leqslant |\mathbf{a}\mathbf{b}|^* \leqslant |\mathbf{a}|^*|\mathbf{b}|^*; \qquad |\mathbf{a}+\mathbf{b}|^* \leqslant |\mathbf{a}|^*+|\mathbf{b}|^*.$$

Quart. J. Math. Oxford (2), 2 (1951), 81–4

The first two lines are clear. To consider the third, put $\mathbf{c} = \mathbf{ab}$, and let $\mathbf{w}$ be a unit latent vector of $\bar{\mathbf{c}}'\mathbf{c}$ corresponding to the characteristic value $|\mathbf{c}|^{*2}$. Write

$$\xi_x = \mathbf{ax}, \qquad \eta_x = \mathbf{bx}, \qquad \zeta_x = \mathbf{cx},$$

where $\mathbf{x}$ is a column vector. Then, by the argument indicated in § 3,

$$\bar{\zeta}'_x \zeta_x \leqslant \bar{\zeta}'_w \zeta_w = |\mathbf{c}|^{*2}, \quad \text{if } \bar{\mathbf{x}}'\mathbf{x} \leqslant 1;$$

$$\bar{\eta}'_x \eta_x \leqslant |\mathbf{b}|^{*2}, \quad \text{if } \bar{\mathbf{x}}'\mathbf{x} \leqslant 1;$$

and, for any real number k,

$$\bar{\xi}'_x \xi_x \leqslant k^2 |\mathbf{a}|^{*2} \quad \text{if } \bar{\mathbf{x}}'\mathbf{x} \leqslant k^2.$$

Now $$\zeta_x = \mathbf{abx} = \mathbf{a}\eta_x = \xi_{\eta_x};$$

and, since we now have

$$\bar{\eta}'_x \eta_x \leqslant |\mathbf{b}|^{*2}, \quad \text{if } \bar{\mathbf{x}}'\mathbf{x} \leqslant 1,$$

and $$\bar{\xi}'_{\eta_x} \xi_{\eta_x} \leqslant |\mathbf{a}|^{*2} |\mathbf{b}|^{*2}, \quad \text{if } \bar{\eta}'_x \eta_x \leqslant |\mathbf{b}|^{*2},$$

it follows that $$\bar{\zeta}'_x \zeta_x \leqslant |\mathbf{a}|^{*2} |\mathbf{b}|^{*2}, \quad \text{if } \bar{\mathbf{x}}'\mathbf{x} \leqslant 1.$$

Hence, taking $\mathbf{x} = \mathbf{w}$, we get $|\mathbf{ab}|^* \leqslant |\mathbf{a}|^* |\mathbf{b}|^*$; and similarly

$$|\mathbf{a}|_* |\mathbf{b}|_* \leqslant |\mathbf{ab}|_*.$$

Finally, put $\mathbf{c} = \mathbf{a}+\mathbf{b}$. Then, if $\bar{\mathbf{x}}'\mathbf{x} \leqslant 1$,

$$\begin{aligned} \bar{\zeta}'_x \zeta_x &= \bar{\xi}'_x \xi_x + \bar{\xi}'_x \eta_x + \bar{\eta}'_x \xi_x + \bar{\eta}'_x \eta_x \\ &\leqslant \bar{\xi}'_x \xi_x + 2(\bar{\xi}'_x \xi_x)^{\frac{1}{2}} (\bar{\eta}'_x \eta_x)^{\frac{1}{2}} + \bar{\eta}'_x \eta_x \\ &\leqslant (|\mathbf{a}|^* + |\mathbf{b}|^*)^2, \end{aligned}$$

by Cauchy's inequality. Hence $|\mathbf{a}+\mathbf{b}|^* \leqslant |\mathbf{a}|^* + |\mathbf{b}|^*$.

4.1. Corollary. *For any matrix sequence* $\mathbf{a}_1, \mathbf{a}_2, \ldots$, *and for any* n, *finite or infinite,*

$$\prod_{\nu=1}^{n} |\mathbf{a}_\nu|_* \leqslant \left| \prod_{\nu=1}^{n} \mathbf{a}_\nu \right|_* \leqslant \left| \prod_{\nu=1}^{n} \mathbf{a}_\nu \right|^* \leqslant \prod_{\nu=1}^{n} |\mathbf{a}_\nu|^*,$$

$$\left| \sum_{\nu=1}^{n} \mathbf{a}_\nu \right|^* \leqslant \sum_{\nu=1}^{n} |\mathbf{a}_\nu|^*.$$

Proof is by induction for every finite n, and then letting $n \to \infty$.

5. Lemma. *A matrix tends to the null matrix as its upper modulus tends to zero:* $\mathbf{a} \to \mathbf{0}$ *as* $|\mathbf{a}|^* \to 0$.

By § 3, $$0 \leqslant \bar{\mathbf{x}}'\bar{\mathbf{a}}'\mathbf{ax} \leqslant |\mathbf{a}|^{*2}$$

when $\bar{\mathbf{x}}'\mathbf{x} = 1$. Thus the field of values for the Hermitian matrix $\bar{\mathbf{a}}'\mathbf{a}$ is bounded by $|\mathbf{a}|^{*2}$, its maximum characteristic value.

Now the νth diagonal element of $\bar{\mathbf{a}}'\mathbf{a}$ is $\bar{\mathbf{a}}'_{(\nu}\mathbf{a}_{\nu)}$, where $\mathbf{a}_{\nu)}$ is the νth column of $\mathbf{a}$ and $\bar{\mathbf{a}}'_{(\nu}$ the νth row of $\bar{\mathbf{a}}'$. It is a value of $\bar{\mathbf{a}}'\mathbf{a}$, attained for

the νth coordinate vector: that is, the vector which has unity for the νth element and all other elements zero.

Hence, for each ν, $\quad 0 \leqslant \bar{\mathbf{a}}'_{(\nu}\,\mathbf{a}_{\nu)} \leqslant |\mathbf{a}|^{*2}$.

In other words $|\mathbf{a}|^{*2}$ bounds the sums of the squares of the elements in any column of $\mathbf{a}$, from which the lemma follows.

6. THEOREM. *For any matrix sequence* $\mathbf{a}_1, \mathbf{a}_2, \ldots$,

(i) $\prod_{\nu=1}^{\infty} \mathbf{a}_\nu = \mathbf{0}$ *if* $\overline{\lim_{\nu\to\infty}} |\mathbf{a}_\nu|^* < 1$;

(ii) $\sum_{\nu=1}^{\infty} \mathbf{a}_\nu$ *is convergent if* $\sum_{\nu=1}^{\infty} |a_\nu|^*$ *is convergent;*

(iii) $\prod_{\nu=1}^{\infty} (1+\mathbf{a}_\nu)$ *is convergent if* $\prod_{\nu=1}^{\infty} (1+|a_\nu|^*)$ *is convergent.*

These are consequences of the corollary of § 4.1 and the lemma of § 5. Thus, if $\overline{\lim_{\nu\to\infty}} |\mathbf{a}_\nu|^* < 1$, then, by the corollary,

$$0 \leqslant \left|\prod_{\nu=1}^{\infty} \mathbf{a}_\nu\right|^* \leqslant \prod_{\nu=1}^{\infty} |\mathbf{a}_\nu|^* = 0,$$

and then, by the lemma, $\quad \prod_{\nu=1}^{\infty} \mathbf{a}_\nu = \mathbf{0}$,

which proves the first result.

For the second result, suppose that $\sum_{\nu=1}^{\infty} |\mathbf{a}_\nu|^*$ is convergent. Then the partial sums form a Cauchy sequence, so that, by use of the corollary in § 4.1,

$$0 \leqslant \left|\sum_{\nu=m+1}^{n} \mathbf{a}_\nu\right|^* \leqslant \sum_{\nu=m+1}^{n} |\mathbf{a}_\nu|^* \to 0 \quad (m, n \to \infty).$$

Then, by the lemma, it follows that

$$\sum_{\nu=m+1}^{n} \mathbf{a}_\nu \to \mathbf{0} \quad (m, n \to \infty);$$

and convergence for the matrix sum follows by the general principle of convergence applied to each of its element sums.

To prove the last part, let

$$\mathbf{P}_n = \prod_{\nu=1}^{n} (1+\mathbf{a}_\nu), \qquad \mathbf{P}^*_n = \prod_{\nu=1}^{n} (1+|\mathbf{a}_\nu|^*);$$

and for $m \leqslant n$, let

$$\mathbf{P}_{mn} = \mathbf{P}_n/\mathbf{P}_m, \qquad P^*_{mn} = P^*_n/P^*_m.$$

Then it follows from the theorem of § 4 that

$$|\mathbf{P}_{mn}-\mathbf{1}|^* \leqslant P^*_{mn}-1;$$

and further, $\qquad \mathbf{P}_n-\mathbf{P}_m = \mathbf{P}_m(\mathbf{P}_{mn}-\mathbf{1}),$

so that $$|\mathbf{P}_n - \mathbf{P}_m|^* \leqslant P_m^*(P_{mn}^* - 1).$$

Now suppose the convergence:

$$P_n^* \to P^* \quad (n \to \infty);$$

so that $$P_{mn}^* \to 1 \quad (m, n \to \infty).$$

Then $$|\mathbf{P}_n - \mathbf{P}_m|^* \to 0 \quad (m, n \to \infty);$$

and hence $$\mathbf{P}_n - \mathbf{P}_m \to \mathbf{0} \quad (m, n \to \infty);$$

and the convergence of P_n $(n \to \infty)$ follows from the general principle of convergence applied to each of the element sequences.

7. THEOREM. *For a matrix function* $\mathbf{f} = \mathbf{f}(\mathbf{a}_1, \mathbf{a}_2, \ldots)$, *formed from sums and products in a set* $\mathbf{a}_1, \mathbf{a}_2, \ldots$ *of matrix arguments, if* $f^* = f^*(\mathbf{a}_1, \mathbf{a}_2, \ldots)$ *is the corresponding scalar function in the upper moduli* $|\mathbf{a}_1|^*, |\mathbf{a}_2|^*, \ldots$ *of the arguments, then* $\mathbf{f}$ *is defined whenever* f^* *is, and, if* ϕ *is any characteristic value of* $\mathbf{f}$, *then* $0 \leqslant |\phi| \leqslant f^*$.

Writing $\mathbf{f}$ as a sum $\sum_\nu \mathbf{p}_\nu$ of products $\mathbf{p}_\nu = \mathbf{p}_\nu(\mathbf{a}_1, \mathbf{a}_2, \ldots)$ in the matrices $\mathbf{a}_1, \mathbf{a}_2, \ldots$, let $p_\nu^* = p_\nu^*(|\mathbf{a}_1|^*, |\mathbf{a}_2|^*, \ldots)$ be the corresponding product in the scalars $|\mathbf{a}_1|^*, |\mathbf{a}_2|^*, \ldots$, and write $f^* = f^*(\mathbf{a}_1, \mathbf{a}_2, \ldots)$ for the scalar sum $\sum_\nu p_\nu^*$.

Then, by § 4.1, $$|\mathbf{f}|^* \leqslant \sum_\nu |p_\nu|^* \leqslant \sum_\nu p_\nu^*,$$

and hence $$|\mathbf{f}|^* \leqslant f^*.$$

Hence, by § 3, for any characteristic value ϕ of $\mathbf{f}$,

$$0 \leqslant |\phi| \leqslant f^*.$$

That a series expressing the function $\mathbf{f}$ is convergent whenever a corresponding one for the function f^* is convergent, is clear from the theorem of § 6.

The origin of this note is in seeking a generalization for the theorem of Oldenberger's (**1**) that the infinite power of a matrix is null if and only if the moduli of its characteristic values are all less than unity. This is partially achieved in the first result of § 6. Criteria for the convergence of certain numerical iterations can be established by using this result.

The theorem of § 3 has been previously proved on different lines by Browne (**2**).

The author's thanks are due to the referee for some helpful criticism.

REFERENCES

1. R. Oldenberger, 'Infinite powers of matrices and characteristic roots', *Duke Math. Journ.* 6 (1940), 257.
2. E. T. Browne, 'The characteristic equation of a matrix', *Bull. American Math. Soc.* 34 (1928), 363–8.

AN ASYMPTOTIC FORMULA IN THE THEORY OF PARTITIONS

By G. SZEKERES (*Adelaide*)

[Received 18 January 1950]

1. Introduction

THE purpose of this paper is to investigate the asymptotic behaviour, as $k \to \infty$, $n \to \infty$, of the following functions in number theory:

$p(n,k)$, the number of partitions of n into exactly k positive integral parts;

$P(n,k) = \sum_{r \leq k} p(n,r)$;

$q(n,k)$, the number of partitions of n into exactly k positive unequal parts.

We do not need to consider each one of these functions separately. It follows from the identities

$$P(n,k) = p(n+k,k) = q\left\{n+\binom{k+1}{2},k\right\} \tag{1.1}$$

that once we have an asymptotic expression for, say, $P(n,k)$, then a trivial change in the variables will immediately furnish an expression for $p(n,k)$ or $q(n,k)$, and conversely. Actually we shall find it more convenient to use a fourth arithmetical function $R(N,k)$ defined by

$$R(N,k) = P\{N-\tfrac{1}{4}k(k+1),k\}$$

whose generating function can be managed more easily. From the asymptotic expression for $R(N,k)$ we obtain the corresponding expressions for $P(n,k)$, $p(n,k)$, $q(n,k)$ by putting

$$N = n+\tfrac{1}{4}k(k+1), \qquad N = n+\tfrac{1}{4}k(k-3), \qquad N = n-\tfrac{1}{4}k(k+1)$$

respectively.

The problem of this paper was considered previously by Erdős and Lehner,* who investigated $P(n,k)$ in the neighbourhood of $k = \frac{1}{2}cn^{\frac{1}{2}}\log n$, $c = 6^{\frac{1}{2}}\pi^{-1}$, which is the 'normal' number of summands in the partitions of n. They found that

$$P(n,k)/P(n) = \{1+o(1)\}\exp(-ce^{-\lambda}) \tag{1.2}$$

* P. Erdős and J. Lehner, *Duke Math. Journ.* 8 (1941), 335–45.

for $k = \frac{1}{2}cn^{\frac{1}{2}}\log n + c\lambda n^{\frac{1}{2}}$, where $P(n) = P(n,n)$ is the number of unrestricted partitions of n. Using their method, Auluck, Chowla, and Gupta* proved

$$n^{\frac{1}{2}}p(n,k)/P(n) = \{1+o(1)\}\exp(-\lambda - ce^{-\lambda})$$

for $k = \frac{1}{2}cn^{\frac{1}{2}}\log n + c\lambda n^{\frac{1}{2}}$. Erdős and Lehner† also investigated $P(n,k)$ for small values of k and proved that

$$P(n,k),\ p(n,k),\ q(n,k) \cong \frac{1}{k!}\binom{n-1}{k-1} \tag{1.3}$$

uniformly for $k = o(n^{\frac{1}{2}})$. This result can be obtained easily from the elementary inequalities

$$\frac{1}{k!}\binom{n-\binom{k}{2}-1}{k-1} \leqslant q(n,k) \leqslant \frac{1}{k!}\binom{n-1}{k-1} \leqslant p(n,k) \leqslant \frac{1}{k!}\binom{n+\binom{k}{2}-1}{k-1}$$

and

$$\frac{1}{k!}\binom{n+k-1}{k-1} \leqslant P(n,k) \leqslant \frac{1}{k!}\binom{n+\binom{k+1}{2}-1}{k-1},$$

which follow from the recursive formula

$$q(n,k) = \sum_{r=1}^{[n/k]} q(n-rk, k-1). \tag{1.4}$$

To prove, for example,

$$q(n,k) \leqslant \frac{1}{k!}\binom{n-1}{k-1},$$

assume induction on k:

$$q(n,k) \leqslant \frac{1}{(k-1)!}\sum_{r>0}\binom{n-rk-1}{k-2} \leqslant \frac{1}{k!}\sum_{r>0}\binom{n-r-1}{k-2} = \frac{1}{k!}\binom{n-1}{k-1}.$$

The first inequality for $q(n,k)$ can be proved similarly, and the inequalities for $p(n,k)$ and $P(n,k)$ follow at once from (1.1).‡

In the present paper I shall obtain an asymptotic formula for $R(N,k)$ which holds in a much wider range than (1.3), namely for $N > Ck^2$, where the constant C has the value 0·385 approximately.§ More precisely, I shall prove the following theorem:

* F. C. Auluck, S. Chowla, and H. Gupta, *J. Indian Math. Soc.* 6 (1942), 105–12.

† P. Erdős and J. Lehner, *Duke Math. Journ.* 8 (1941), 335–45.

‡ Essentially the same proof was given by H. Gupta for $p(n,k)$, *Proc. Indian Acad. Sci.* (A) 16 (1942), 101–2, and by F. C. Auluck for $P(n,k)$, *J. Indian Math. Soc.* 6 (1942), 113–14.

§ The best possible value for C is 0·25, since $N \geqslant \frac{1}{4}k(k+1)$.

THEOREM 1. *Let*

$$u = u(t) = t + \tfrac{1}{36}t^3 + \tfrac{41}{32400}t^5 + \dots \quad (|t| < 4)$$

be defined as the inverse function of

$$t = u^2\left(\int_0^u \tfrac{1}{2}x \coth \tfrac{1}{2}x\, dx\right)^{-1} = \left(\tfrac{1}{4} + u^{-2}\int_0^u \frac{x}{e^x - 1}\, dx\right)^{-1}, \qquad (1.5)$$

and let $\rho = 2{\cdot}598\dots$ *be the radius of convergence of* $v = v(t)$ *defined by*

$$t = v^2\left(\int_0^v (2 - \tfrac{1}{2}x \cot \tfrac{1}{2}x)\, dx\right)^{-1}.$$

Write $\alpha = (k+\frac{1}{2})^2/N$. *Then, for every* $m > 0$,

$$\log R(N,k) = k\left\{2\frac{u(\alpha)}{\alpha} - \log\{2\sinh \tfrac{1}{2}u(\alpha)\}\right\} - \log N + \\ + \sum_{\mu=0}^{m-1} k^{-\mu}\psi_\mu(\alpha) + O(k^{-m}) \qquad (1.6)$$

uniformly for $\alpha \leqslant \rho_0 < \rho$, *where*

$$\psi_0(t) = \frac{u(t)}{t} - \tfrac{1}{2}\log\left(\frac{2}{u(t)}\sinh \tfrac{1}{2}u(t)\right) + \tfrac{1}{2}\log u'(t) - \log(2\pi) \qquad (1.6')$$

and the functions $\psi_\mu(t)$ *are analytic for* $|t| < \rho$.

Although I can prove (1.6) only for $\alpha \leqslant \rho_0 < 2{\cdot}598\dots$, and hence for $N = (k+\frac{1}{2})^2/\alpha \geqslant 0{\cdot}385k^2$, it seems very likely that it holds uniformly for every $N \geqslant Ck^2$ with $C > \frac{1}{4}$. The following argument suggests that (1.6) is valid even in the critical neighbourhood of $C = \frac{1}{4}$ (or $\alpha = 4$), at least for $m = 1$ if $O(k^{-1})$ is replaced by $o(1)$. For, let us make this assumption and take the function $P(n,k)$ at $k = n$. We have

$$\alpha = (k+\tfrac{1}{2})^2/N = (k+\tfrac{1}{2})^2\{k + \tfrac{1}{4}k(k+1)\}^{-1};$$

hence, from (1.5) at $t = \alpha$,

$$u^{-2}\int_0^u \frac{x}{e^x - 1}\, dx = \frac{1}{\alpha} - \frac{1}{4} = (k - \tfrac{1}{16})(k+\tfrac{1}{2})^{-2},$$

$$u = c^{-1}(k+\tfrac{1}{2})(k-\tfrac{1}{16})^{-\frac{1}{2}}\{(1 + o(k^{-2})\},$$

and

$$2\frac{u}{\alpha} - \log(2\sinh \tfrac{1}{2}u) = 2u\left(\frac{1}{\alpha} - \frac{1}{4}\right) - \log(1 - e^{-u}) \\ = 2c^{-1}(k+\tfrac{1}{2})^{-1}(k-\tfrac{1}{16})^{\frac{1}{2}}\{1 + o(k^{-2})\}.$$

Also, from (1.5),

$$\frac{du}{dt} = ut^{-2}\left(\frac{2}{t} - \tfrac{1}{2}\coth \tfrac{1}{2}u\right)^{-1} = ut^{-2}\left\{2\left(\frac{1}{t} - \frac{1}{4}\right) - \frac{1}{e^u - 1}\right\}^{-1}, \qquad (1.7)$$

and hence $$u'(\alpha) = \tfrac{1}{32}c^{-1}k^{\frac{1}{2}}\{1+o(k^{-1})\}.$$

Putting these in (1.6), (1.6′), we easily obtain

$$P(n,n) = \{1+o(1)\}4^{-1}3^{-\frac{1}{2}}n^{-1}\exp(2c^{-1}n^{\frac{1}{2}}),$$

which is the asymptotic term of the Hardy–Ramanujan formula. By a somewhat lengthier computation we can also derive (1.2) from the formula if we assume its validity in the region $\alpha = 4-\delta/\log^2 k$, $\delta > 0$.

The formula (1.6) remains valid if $\alpha \to 0$ as $k \to \infty$: that is, if $k = o(n^{\frac{1}{2}})$. We then have $$u(\alpha)/\alpha \to 1, \quad u'(\alpha) \to 1,$$

and

$$\log R(N,k) = k\left(2u(\alpha)/\alpha - \log\left(\frac{2}{\alpha}\sinh\tfrac{1}{2}u(\alpha)\right)\right) - \log(2\pi N\alpha^k) + 1 + o(1)$$

$$= k(2-\tfrac{1}{72}\alpha^2+\tfrac{7}{10800}\alpha^4-\ldots)+(k-1)\log N - 2k\log k - \log(2\pi)+o(1).$$

Notably, if $k = O(n^{\frac{1}{2}-\delta})$, $\delta > 0$, then, for $q(n,k)$,

$$(k-1)\log N = (k-1)\log n + (k-1)\log\{1-\tfrac{1}{4}k(k+1)/n\}$$

$$= (k-1)\log n - (\tfrac{1}{4}k^3/n+\tfrac{1}{32}k^5/n^2+\tfrac{1}{192}k^7/n^3+\ldots)+o(1).$$

Hence

$$q(n,k) \cong n^{k-1}\{k!(k-1)!\}^{-1}\exp\{-(\tfrac{1}{4}k^3/n+\tfrac{13}{288}k^5/n^2+\tfrac{1}{192}k^7/n^3+\ldots)\}$$

by Stirling's formula. Similarly

$$p(n,k),\ P(n,k) \cong n^{k-1}\{k!(k-1)!\}^{-1}\exp(\tfrac{1}{4}k^3/n-\tfrac{13}{288}k^5/n^2+\tfrac{1}{192}k^7/n^3-\ldots).$$

Finally, if $k = O(n^{\frac{1}{2}})$, then

$$q(n,k) \cong n^{k-1}\{k!(k-1)!\}^{-1}\exp(-\tfrac{1}{4}k^3/n)$$

and $$p(n,k),\ P(n,k) \cong n^{k-1}\{k!(k-1)!\}^{-1}\exp(\tfrac{1}{4}k^3/n).$$

Generally, it follows from (1.6) that $p(n,k)$ and $P(n,k)$ are asymptotically equal if $k = o(n^{\frac{1}{2}})$. This was proved by Erdős and Lehner for $k = o(n^{\frac{1}{2}})$ and conjectured for $k = o(n^{\frac{1}{2}})$.

2. A conjecture of Auluck, Chowla, and Gupta

For a fixed n, $p(n,k)$ has a maximum at a certain $k = k_0(n)$. Erdős proved* that

$$k_0 = \tfrac{1}{2}cn^{\frac{1}{2}}\log n + c\log cn^{\frac{1}{2}} + o(n^{\frac{1}{2}}) \quad (c = 6^{\frac{1}{2}}/\pi). \tag{2.1}$$

He also stated that, if $q(n,k)$ has its maximum (for a fixed n) at k_0, then

$$k_0 = (2^{\frac{1}{2}}\log 2)cn^{\frac{1}{2}} + dn^{\frac{1}{2}} + o(n^{\frac{1}{2}}) \tag{2.2}$$

for a certain constant d which he did not specify. Auluck, Chowla, and

* P. Erdős, *Bull. American Math. Soc.* 52 (1946), 185–8.

Gupta conjectured* that $p(n,k)$ is a *monotonous* function of k: that is, $p(n,k_1) < p(n,k_2)$ if $k_1 < k_2 < k_0$ and $p(n,k_1) > p(n,k_2)$ if $k_0 < k_1 < k_2$. I shall now show that the analogous conjecture for $q(n,k)$ is true if n is sufficiently large and $0{\cdot}635k^2 \leqslant n$.†

For a fixed n,

$$\alpha = \alpha_k = (k+\tfrac{1}{2})^2\{n+\tfrac{1}{16}-\tfrac{1}{4}(k+\tfrac{1}{2})^2\}^{-1}$$

is a function of k. We have

$$\Delta\alpha_k = \alpha_k - \alpha_{k-1} = \tfrac{1}{2}\beta(\beta+4)k^{-1} + O(k^{-3}), \tag{2.3}$$

where $$\beta = k^2(n+\tfrac{1}{16}-\tfrac{1}{4}k^2)^{-1}.$$

Also, if $\phi(t)$ is analytic at $t = \beta$, then

$$\Delta\phi(\alpha_k) = \phi(\alpha_k) - \phi(\alpha_{k-1}) = \phi'(\beta)\Delta\alpha_k + O\{(\Delta\alpha_k)^3\}$$
$$= \beta(\beta+4)\phi'(\beta)/2k + O(k^{-3})$$

and $$\phi(\alpha_k) + \phi(\alpha_{k-1}) = 2\phi(\beta) + O(k^{-2}).$$

From (1.6) and (1.6′) we have, if

$$\alpha = N^{-1}(k+\tfrac{1}{2})^2 = \{n-\tfrac{1}{4}k(k+1)\}^{-1}(k+\tfrac{1}{2})^2 \leqslant \rho_0 < 2{\cdot}598\ldots,\ n \geqslant 0{\cdot}635k^2,$$

$$\log q(n,k) = (k+\tfrac{1}{2})[2u(\alpha_k)/\alpha_k - \log\{2\sinh\tfrac{1}{2}u(\alpha_k)\}] +$$
$$+\tfrac{1}{2}\log\{uu'(\alpha_k)\} + \log\alpha_k - \log(k+\tfrac{1}{2})^2 - \log(2\pi) + k^{-1}\psi_1(\alpha_k) + O(k^{-2}),$$

$$\log q(n,k) - \log q(n,k-1)$$
$$= k\left(2\frac{u'}{\beta} - 2\frac{u}{\beta^2} - \frac{u'}{2}\coth\tfrac{1}{2}u\right)\Delta\alpha_k + 2\frac{u}{\beta} - \log(2\sinh\tfrac{1}{2}u) +$$
$$+\frac{1}{2}\left(\frac{u'}{u}+\frac{u''}{u'}\right)\Delta\alpha_k + \frac{1}{\beta}\Delta\alpha_k - \frac{2}{k} + \frac{1}{k}\psi_1'(\beta)\Delta\alpha_k + O(k^{-2})$$
$$= \beta(\beta+4)\left(\frac{u'}{\beta} - \frac{u}{\beta^2} - \frac{u'}{4}\coth\tfrac{1}{2}u\right) + 2\frac{u}{\beta} - \log(2\sinh\tfrac{1}{2}u) +$$
$$+\left(\frac{u'}{u}+\frac{u''}{u'}\right)\frac{\beta(\beta+4)}{4k} + \frac{\beta+4}{2k} - \frac{2}{k} + O(k^{-2}),$$

where u, u', u'' are to be taken at the place β. If we introduce the expression (1.7) for u', we obtain

$$\Delta\log q(n,k) = -\log(e^u-1) + \frac{\beta(\beta+4)}{4k}\left(\frac{u'}{u}+\frac{u''}{u'}\right) + \frac{\beta}{2k} + O(k^{-2}). \tag{2.4}$$

* F. C. Auluck, S. Chowla, and H. Gupta, *J. Indian Math. Soc.* 6 (1942), 105–12.

† These restrictions on n and k are not really important. The first condition (that n be large) can probably be discarded if we replace the O-notations throughout the proof of Theorem 1 by explicit constants and check the conjecture for smaller values of n. But, if we know that the conjecture is true for every n, k satisfying, say, $0{\cdot}9k^2 \leqslant n$, then its validity for $k^2 > n$ can easily be established by induction on n, using the recursion (1.4).

This cannot be zero unless

$$\log\{e^{u(\beta)}-1\} = O(k^{-1}),$$

i.e.

$$u(\beta) = \log 2+O(k^{-1}),$$

$$\frac{1}{\beta}-\frac{1}{4} = u^{-2}\int_0^u \frac{x}{e^x-1}\,dx = \tfrac{1}{2}\{(c\log 2)^{-2}-1\}+O(k^{-1}),$$

$$\beta = \{(2c^2\log^2 2)^{-1}-\tfrac{1}{4}\}^{-1}+O(k^{-1}),$$

since $$\int_0^{\log 2} \frac{x}{e^x-1}\,dx = \tfrac{1}{2}(\tfrac{1}{6}\pi^2-\log^2 2) = \tfrac{1}{2}(c^{-2}-\log^2 2).*$$

Note that $\{(2c^2\log^2 2)^{-1}-\frac{1}{4}\}^{-1} < 0{\cdot}7$, and hence β is well within the region of validity of (1.6).

From (1.7) we have

$$\frac{u'}{u}(t)+\frac{u''}{u'}(t) = \left\{2(1-\tfrac{1}{4}t)-\frac{t}{e^u-1}\right\}^{-1}\left\{1+\frac{2}{e^u-1}-\frac{te^u}{(e^u-1)^2}u'\right\}.$$

Hence for $t=\beta$,

$$u' = \log 2\{2\beta(1-\tfrac{3}{4}\beta)\}^{-1}+O(k^{-1}),$$

$$\frac{u'}{u}+\frac{u''}{u'} = (2-\tfrac{3}{2}\beta)^{-1}\{3-(1-\tfrac{3}{4}\beta)^{-1}\log 2\}+O(k^{-1}).$$

Therefore the right-hand side of (2.4) cannot be zero unless

$$u(\beta) = \log\left(2+\frac{b}{k}\right)+O(k^{-2})$$

with

$$b = \tfrac{1}{4}\beta(\beta+4)\left(\frac{u'}{u}+\frac{u''}{u'}\right)+\tfrac{1}{2}\beta+O(k^{-1}) = \frac{2}{a-1}-\frac{a\log 2}{2(a-1)^2},$$

$$a = \tfrac{1}{12}(\pi/\log 2)^2. \qquad (2.5)$$

If $u(\beta) < \log(2+b/k)$, the right-hand side of (2.4) is certainly positive unless

$$u(\beta) = \log(2+b/k)+O(k^{-2}),$$

i.e.

$$\beta^{-1} = \tfrac{1}{4}+\left(\log 2+\frac{b}{2k}\right)^{-2}\int_0^{\log 2+b/2k} \frac{x}{e^x-1}\,dx+O(k^{-2})$$

$$= a-\tfrac{1}{4}-b(a-1)(k\log 2)^{-1}+O(k^{-2}). \qquad (2.6)$$

Similarly (2.4) is negative for $u(\beta) > \log(2+b/k)$ unless (2.6) holds.

* See D. Bierens de Haan, *Nouvelles Tables d'Intégrales définies* (New York, 1939), 151, formula (104, 5).

Suppose now that β satisfies (2.6) and let k_0 be defined by

$$\beta = (k_0+\tfrac{1}{2})^2\{n+\tfrac{1}{16}-\tfrac{1}{4}(k_0+\tfrac{1}{2})^2\}^{-1}. \tag{2.7}$$

Then, if $k_1 \leqslant k_0$ and $\alpha_1 = k_1^2(n+\frac{1}{16}-\frac{1}{4}k_1^2)^{-1}$, we have

$$(\alpha_1)^{-1} \geqslant k_0^{-2}(n+\tfrac{1}{16}-\tfrac{1}{4}k_0^2) \geqslant (1+1/k_0)/\beta,$$

which shows that $(\alpha_1)^{-1}$ does not satisfy (2.6) if k_0 is large. Hence $\Delta \log q(n, k_1) > 0$. Similarly we can show that, if $k_2 \geqslant k_0+1$, then $\Delta \log q(n, k_2) < 0$. This proves the conjecture for large values of $n \geqslant 0{\cdot}635k^2$.

From (2.6) and (2.7) we obtain for the maximum k_0

$$(n+\tfrac{1}{16})(k_0+\tfrac{1}{2})^{-2}-\tfrac{1}{4} = a-\tfrac{1}{4}-b(a-1)\{(k_0+\tfrac{1}{2})\log 2\}^{-1}+O(k_0^{-2}),$$

$$n = a(k_0+\tfrac{1}{2})^2-b(a-1)(k_0+\tfrac{1}{2})/\log 2+O(1),$$

$$k_0 = a^{-\frac{1}{2}}n^{\frac{1}{2}}+b\left(1-\frac{1}{a}\right)\Big/2\log 2-\tfrac{1}{2}+O(n^{-\frac{1}{2}})$$

$$= 0{\cdot}764304...n^{\frac{1}{2}}-0{\cdot}008428...+O(n^{-\frac{1}{2}}). \tag{2.8}$$

This improves the result of Erdős and shows that the constant d in (2.2) is zero. The error term in (2.8) tends to zero. Hence the position of the maximum is exactly known for large values of n.

It is interesting to compare the value of k_0 calculated from (2.8) with data obtained from a table of Todd's* containing all the values of $p(n, k)$ for $n \leqslant 100$. We obtain for $n = 100$, calculated from the table by means of $q(n, k) = p\left\{n-\binom{k}{2}, k\right\}$,

$$q(100, 6) = 65827, \qquad q(100, 7) = 108869,$$

$$q(100, 8) = 116263, \qquad q(100, 9) = 79403,$$

whereas (2.8) gives $k_0 \sim 7{\cdot}63$. But even for small values of n the formula gives the correct value of k_0. For $n = 16$,

$$q(16, 2) = 7, \qquad q(16, 3) = 14, \qquad q(16, 4) = 9,$$

whereas (2.8) gives $k_0 \sim 3{\cdot}05$.

Another result which follows easily from my formula is an asymptotic expression for $Q(n)$, the total number of partitions of n into unequal parts. The function

$$Q(n, k) = \sum_{r<k} q(n, r)$$

is not connected in a simple manner with $R(N, k)$; we therefore have to use the formula

$$Q(n) = \sum_k q(n, k). \tag{2.9}$$

* J. A. Todd, *Proc. London Math. Soc.* (2) 48 (1945), 229–42.

Let us first consider those terms which belong to k's in the neighbourhood of k_0.

Write $u(\alpha_k) = \{1+\lambda(k)\}\log 2$, $\lambda = \lambda(k) = O(n^{-\frac{1}{4}})$, and put

$$a = \tfrac{1}{12}(\pi/\log 2)^2 = 1{\cdot}71\ldots.$$

By an easy calculation

$$\frac{1}{\alpha} = \frac{1}{\alpha_k} = \frac{1}{4} + u^{-2}\int_0^u \frac{x}{e^x-1}\,dx$$

$$= a - \tfrac{1}{2} - 2(a-1)\lambda + (3a-3-\log 2)\lambda^2 + O(\lambda^3),$$

$$k+\tfrac{1}{2} = (n+\tfrac{1}{16})^{\frac{1}{4}}\left(\frac{1}{\alpha}+\frac{1}{4}\right)^{-\frac{1}{4}}$$

$$= (n+\tfrac{1}{16})^{\frac{1}{4}}a^{-\frac{1}{4}}\left\{1+\left(1-\frac{1}{a}\right)\lambda+\frac{1}{2a}\left(\frac{3}{a}-3+\log 2\right)\lambda^2+O(\lambda^3)\right\},$$

$$2u(\alpha)/\alpha - \log\{2\sinh\tfrac{1}{2}u(\alpha)\}$$

$$= 2\log 2\{a-(a-1)\lambda+(a-1-\tfrac{1}{2}\log 2)\lambda^2+O(\lambda^3)\},$$

$$(k+\tfrac{1}{2})\left(\frac{2u}{\alpha}-\log(2\sinh\tfrac{1}{2}u)\right) = 3^{-\frac{1}{4}}\pi(n+\tfrac{1}{16})^{\frac{1}{4}}\left\{1-\frac{1}{2a}\left(1-\frac{1}{a}\right)\lambda^2+O(\lambda^3)\right\},$$

$$\{uu'(\alpha)\}^{\frac{1}{4}} = (a-\tfrac{1}{4})\{2(a-1)\}^{-\frac{1}{4}}\log 2 + O(\lambda),$$

$$N = (n+\tfrac{1}{16})\left(1-\frac{1}{4a}\right)\{1+O(\lambda)\}. \tag{2.10}$$

Hence

$$q(n,k) = \frac{1}{2\pi N}(uu')^{-\frac{1}{4}}\exp\left\{(k+\tfrac{1}{2})\left(2\frac{u}{\alpha}-\log(2\sinh\tfrac{1}{2}u)\right)\right\}\{1+O(k^{-1})\}$$

$$= \frac{a\log 2}{2\pi n}\{2(a-1)\}^{-\frac{1}{4}}\exp\left\{3^{-\frac{1}{4}}\pi n^{\frac{1}{4}}\left(1-\frac{1}{2a}\left(1-\frac{1}{a}\right)\lambda^2\right)\right\}\{1+o(1)\}.$$

The condition $\lambda = O(n^{-\frac{1}{4}})$ is certainly satisfied if $k_1 < k < k_2$, where $k_1 = a^{-\frac{1}{4}}n^{\frac{1}{4}} - n^{\frac{1}{16}}$, $k_2 = a^{-\frac{1}{4}}n^{\frac{1}{4}} + n^{\frac{1}{16}}$. Therefore

$$\sum_{k_1<k<k_2} q(n,k) = \{1+o(1)\}\frac{a\log 2}{2\pi n}\{2(a-1)\}^{-\frac{1}{4}}\exp(3^{-\frac{1}{4}}\pi n^{\frac{1}{4}})\times$$

$$\times \sum_{k_1<k<k_2}\exp\left[-3^{-\frac{1}{4}}\pi n^{\frac{1}{4}}\frac{1}{2a}\left(1-\frac{1}{a}\right)\{\lambda(k)\}^2\right]$$

$$= \{1+o(1)\}\frac{a\log 2}{2\pi n}\{2(a-1)\}^{-\frac{1}{4}}a^{-\frac{1}{4}}\left(1-\frac{1}{a}\right)n^{\frac{1}{4}}\times$$

$$\times \int_{\lambda(k_1)}^{\lambda(k_2)}\exp\left\{-3^{-\frac{1}{4}}\pi n^{\frac{1}{4}}\frac{1}{2a}\left(1-\frac{1}{a}\right)\lambda^2\right\}d\lambda,$$

since $$\Delta k = (k+\tfrac{1}{2})-(k-\tfrac{1}{2}) = \{1+o(1)\}a^{-\frac{1}{2}}\left(1-\frac{1}{a}\right)n^{\frac{1}{2}}\,d\lambda,$$

by (2.10). Writing

$$y = \left\{3^{-\frac{1}{2}}\pi n^{\frac{1}{2}}\frac{1}{2a}\left(1-\frac{1}{a}\right)\right\}^{\frac{1}{2}}\lambda,$$

and, noting that, for $k = k_1$ and $k = k_2$, $|\lambda(k)| > C_1 n^{-\frac{1}{4}}$ for a certain positive constant C_1, and hence that $n^{\frac{1}{2}}\{\lambda(k_i)\}^2 \to \infty$ $(i = 1, 2)$, we obtain

$$\sum_{k_1<k<k_2} q(n,k) = \{1+o(1)\}\tfrac{1}{2}a^{\frac{1}{2}}\log 2\pi^{-\frac{1}{2}}n^{-\frac{1}{4}}3^{\frac{1}{4}}\int_{-\infty}^{\infty} e^{-y^2}\,dy\,\exp(3^{-\frac{1}{2}}\pi n^{\frac{1}{2}})$$

$$= \{1+o(1)\}4^{-1}3^{-\frac{1}{4}}n^{-\frac{1}{4}}\exp(3^{-\frac{1}{2}}\pi n^{\frac{1}{2}}).$$

Now it is easily seen that the terms with $k \leqslant k_1$ and $k \geqslant k_2$ in (2.9) do not contribute essentially to the value of $Q(n)$, for $n^{\frac{1}{2}}\lambda^2(k_i) > C_1^2 n^{1/10}$ where $i = 1, 2$. Hence

$$\exp\left[-3^{-\frac{1}{2}}\pi n^{\frac{1}{2}}\frac{1}{2a}\left(1-\frac{1}{a}\right)\{\lambda(k_i)\}^2\right] < \exp(-C_2 n^{1/10})$$

for a positive constant C_2, and

$$q(n,k_i) < q(n,k_0)\exp(-C_2 n^{1/10}),$$

$$\sum_{k<k_1} q(n,k)+\sum_{k>k_2} q(n,k) < n^{\frac{1}{2}}\{q(n,k_1)+q(n,k_2)\} = o\{q(n,k_0)\},$$

which implies $$Q(n) \cong 4^{-1}3^{-\frac{1}{4}}n^{-\frac{1}{4}}\exp(3^{-\frac{1}{2}}\pi n^{\frac{1}{2}}).$$

This is the asymptotic term of the Hardy–Ramanujan formula for $Q(n)$. The above proof is entirely independent of the theorem of residues and the transformation properties of elliptic modular functions, which seem to be indispensable for the proof of the exact formula of Hardy–Ramanujan–Hua.*

For $q(n,k_0) = \max q(n,k)$ we obtain

$$q(n,k_0) \cong \tfrac{1}{4}6^{-\frac{1}{2}}\left(1-\frac{1}{a}\right)^{-\frac{1}{4}}n^{-1}\exp(3^{-\frac{1}{2}}\pi n^{\frac{1}{2}}).$$

Hence $$Q(n)/n^{\frac{1}{2}}q(n,k_0) \to 3^{\frac{1}{4}}2^{\frac{1}{2}}(1-1/a)^{\frac{1}{4}} = 1{\cdot}200\ldots.$$

Except for the explicit value of the constant on the right, this result was also found by Erdős.†

If we assume the validity of (1.6), (1.6′) at $\alpha = 4-\delta/\log^2 k$ $(\delta > 0)$

* L. K. Hua, *Trans. American Math. Soc.* 51 (1942), 194–201. There is a misprint in the main formula on p. 195 of the paper. $\frac{3}{8}(n+\frac{1}{24})$ should read $\frac{1}{3}(n+\frac{1}{24})$.

† P. Erdős, *Bull. American Math. Soc.* 52 (1946), 185–8.

and do not care about the error terms, then also (2.1) can be obtained from the formula. It seems likely that

$$k_0 = \tfrac{1}{2}cn^{\frac{1}{4}}\log n + c\log cn^{\frac{1}{4}} + d + o(1) \sim 0{\cdot}38985\ldots n^{\frac{1}{4}}\log n - 0{\cdot}19403n^{\frac{1}{4}} + d$$

for a certain constant d. Data taken from Todd's table suggests that the value of d must be near to 1·7.

3. Proof of the main theorem

The proof of Theorem 1 is based on two results. One is the classical theory of 'waves' of Sylvester.* Let ζ denote a primitive qth root of unity ($q \leqslant k$) and write $P_\zeta(n,k)$ for the coefficient of t^{-1} in

$$f(\zeta,t) = \zeta^{-n}e^{nt}\prod_{j=1}^{k}(1-\zeta^j e^{-jt})^{-1}. \tag{3.1}$$

Then
$$P(n,k) = \sum_{q=1}^{k}\sum_{\zeta} P_\zeta(n,k) = P_1(n,k) + \sum_{q=2}^{k}\sum_{\zeta} P_\zeta(n,k), \tag{3.2}$$

where the summation is to be taken for every primitive qth root of unity ($q \leqslant k$). This remarkable formula splits $P(n,k)$ into two parts: a principal term $P_1(n,k)$ which governs the asymptotic behaviour of $P(n,k)$, and a circulatory part $\sum'_{q,\zeta} P_\zeta(n,k)$ which depends on the arithmetical properties of n.

It may seem surprising that the Sylvester formula has never been used for the purposes of an asymptotic evaluation. Hardy and Ramanujan, in their famous paper on partitions,† came to a rather unfavourable conclusion with regard to this possibility. I quote from their footnote on p. 76: 'These (cf. Sylvester's) formulae do, of course, furnish incidentally asymptotic formulae for the functions in question. But they are, from this point of view, of a very trivial character: the interest which they possess is algebraical.' The asymptotic formulae to which Hardy and Ramanujan refer here are those which hold for fixed k when n tends to infinity: $p(n,k) \cong n^{k-1}/k!\,(k-1)!$, which incidentally follows easily from the elementary formula (1.3). But there is really no need to confine ourselves to that trivial situation (when k is fixed), since Sylvester's theorem expresses an identity which holds for every n and k. As a matter of fact, the Hardy–Ramanujan formula itself is composed of expressions which strongly suggest a close relationship to the corresponding Sylvester waves for $P(n,n)$.

* See e.g. L. E. Dickson, *History of the theory of numbers* (Washington, 1920), 2, 119.

† G. H. Hardy and S. Ramanujan, *Proc. London Math. Soc.* (2) 17 (1918), 75–115.

The other theorem on which my proof is based is the following asymptotic expansion which I obtained quite recently* for the coefficients of certain power series.

THEOREM 2. *Let*

$$f(z) = \sum_{\nu=1}^{\infty} \frac{1}{\nu} c_\nu z^\nu$$

be a given power series with positive radius of convergence and constant term zero, h an integer, α a complex parameter, σ, τ real constants $(0 < \sigma \leqslant 1)$. *Write*

$$G(z) = e^{K\{z - f(\alpha z)\}} = \sum_{\mu=0}^{\infty} \frac{1}{\mu!} A_\mu(h, \alpha) z^\mu, \qquad K = \frac{1}{\sigma}(h+\tau). \tag{3.3}$$

Let $u = u(t) = u(\sigma, t)$ denote the inverse function of

$$t = \frac{u}{\sigma + u f'(u)}, \qquad u = t\Big(\sigma + \sum_{\nu=1}^{\infty} c_\nu u^\nu\Big), \tag{3.4}$$

and let $v = v(\sigma, t)$ be defined by

$$v = t\Big(\sigma + \sum_{\nu=1}^{\infty} |c_\nu| v^\nu\Big).$$

Then

$$\log A_{h-1} = \log A_{h-1}(h, \alpha) = (h-1)\log K - K \int_0^\alpha \frac{u(t) - \sigma t}{t^2}\, dt + \tau \log \frac{u(\alpha)}{\sigma\alpha} +$$

$$+ \tfrac{1}{2} \log \frac{u'(\alpha)}{\sigma} + \sum_{\mu=1}^{m-1} h^{-\mu} \phi_\mu(\alpha) + O(h^{-m}) \tag{3.5}$$

for every $m > 0$ and certain functions $\phi_\mu(\alpha)$, analytic for $|\alpha| < \rho$, where ρ denotes the radius of convergence of $v(t)$. The expansion (3.5) is uniformly valid for $|\alpha| \leqslant \rho_0 < \rho$.

THEOREM 3. *Suppose that the coefficients of $f(z)$ are not constants but have an asymptotic expansion of the form*

$$c_\nu = C_\nu \Big\{ 1 + \sum_{\mu=1}^{m-1} \binom{\nu+1}{2\mu} E_\mu k^{-2\mu} + \delta_\nu E_m \binom{\nu+1}{2m} k^{-2m} \Big(1 + \frac{a}{k}\Big)^\nu \Big\} \tag{3.6}$$

where C_ν, E_μ, a $(\geqslant 0)$ are constants (independent of k) and $|\delta_\nu| < 1$. Then Theorem 2 still holds if we form $u(t)$ with coefficients C_ν independent of k viz.

$$u = t\Big(\sigma + \sum_{\nu=1}^{\infty} C_\nu u^\nu\Big).$$

* G. Szekeres, 'The asymptotic behaviour of the coefficients of certain power series', *Acta Sci. Math. Szeged*, 12 (1950), 187–98, Theorems 1, 2.

COROLLARY. *If the coefficients of $f(z)$ have the form*

$$c_\nu = C_\nu\{1+\delta_\nu(\nu+1)K^{-1}\} \quad (|\delta_\nu| < 1,\ C_\nu \text{ constant}),$$

and

$$u = t\Big(\sigma + \sum_{\nu=1}^{\infty} C_\nu u^\nu\Big),$$

then

$$\log A_{h-1} = (h-1)\log K - K\int_0^\alpha \frac{u(t)-\sigma t}{t^2}\,dt + O(1). \tag{3.7}$$

I shall obtain the proof of Theorem 1 by applying the asymptotic expansions of Theorems 2, 3 to the coefficients of $f(\zeta, t)$ in (3.1). Write

$$R_\zeta(N,k) = P_\zeta\{N-\tfrac{1}{4}k(k+1), k\},$$

so that $R_\zeta(N,k)$ is the coefficient of t^{-1} in

$$f(\zeta,t) = \zeta^{-N+\frac{1}{4}k(k+1)}e^{Nt}\prod_{j=1}^{k}(e^{\frac{1}{2}jt}-\zeta^j e^{-\frac{1}{2}jt})^{-1}. \tag{3.8}$$

Sylvester's theorem states that

$$R(N,k) = R_1(N,k) + \sum_{q=2}^{k}\sum_{\zeta} R_\zeta(N,k). \tag{3.9}$$

4. The principal term $R_1(N,k)$

I shall prove in this section that $R_1(N,k)$ satisfies the asymptotic expansion (1.6), (1.6′). From (3.8),

$$f(1,t) = \frac{1}{k!}e^{Nt}t^{-k}\prod_{\nu=1}^{k}\tfrac{1}{2}\nu t(\sinh\tfrac{1}{2}\nu t)^{-1}. \tag{4.1}$$

Hence, writing

$$g(1,t) = \exp\Big[Nt - \sum_{\nu=1}^{k}\log\{(\tfrac{1}{2}\nu t)^{-1}\sinh\tfrac{1}{2}\nu t\}\Big] = \sum_{\nu=0}^{\infty}\frac{1}{\nu!}b_\nu(N,k)t^\nu, \tag{4.2}$$

we have

$$R_1(N,k) = \{k!(k-1)!\}^{-1}b_{k-1}(N,k). \tag{4.3}$$

From (4.2),

$$g(1,t) = \exp\Big\{Nt - \sum_{\nu=1}^{k}\sum_{p=1}^{\infty}(-1)^{p-1}\frac{B_p}{2p(2p!)}(\nu t)^{2p}\Big\} \tag{4.4}$$

$$= \exp\Big\{Nt - \sum_{p=1}^{\infty}\Big(\sum_{\nu=1}^{k}\nu^{2p}\Big)(-1)^{p-1}\frac{B_p}{2p(2p)!}t^{2p}\Big\}$$

where B_p denotes the pth Bernoulli number ($B_0 = 1, B_1 = \frac{1}{6}, B_2 = \frac{1}{30}, \ldots$). To evaluate $\sum_{\nu=1}^{k}\nu^{2p}$, we use a modified form of the Euler–Bernoulli formula. Let the numbers D_{2s} be defined by

$$D_0 = 1, \qquad \sum_{\mu=0}^{s}\binom{2s+1}{2\mu}D_{2\mu} = 0 \quad (s = 1, 2, \ldots). \tag{4.5}$$

[See Nörlund,* where it is shown that $D_{2s} = (-1)^s(2^{2s}-2)B_s$.]

* N. E. Nörlund, *Differenzenrechnung* (Berlin, 1924), 27–8.

Then, for $m > 0$, $p > 0$,

$$(2p+1)\sum_{\nu=1}^{k}\nu^{2p}$$

$$= \sum_{\mu=0}^{m-1}\binom{2p+1}{2\mu}2^{-2\mu}D_{2\mu}(k+\tfrac{1}{2})^{2p+1-2\mu}+\delta B_m\binom{2p+1}{2m}(k+1)^{2p-2m+1}$$

$$(|\delta| < 1). \quad (4.6)$$

The formula is a direct consequence of a formula given by Nörlund [loc. cit. 31].

Combining (4.6) and (4.4), we obtain

$$g(1,t) = \exp\Big\{Nt-$$

$$-\sum_{\mu=0}^{m-1}2^{-2\mu}D_{2\mu}\sum_{p=1}^{\infty}(-1)^{p-1}\frac{B_p}{2p(2p+1)!}\binom{2p+1}{2\mu}(k+\tfrac{1}{2})^{2p-2\mu+1}t^{2p}-$$

$$-B_m\sum_{p=1}^{\infty}(-1)^{p-1}\frac{B_p}{2p(2p+1)!}\binom{2p+1}{2m}(k+1)^{2p-2m+1}t^{2p}\Big\} \quad (|\delta_p| < 1)$$

or equivalently, introducing

$$z = \frac{N}{k+\frac{1}{2}}t = \frac{1}{\alpha}(k+\tfrac{1}{2})t, \qquad \alpha = (k+\tfrac{1}{2})^2/N,$$

$$G(1,z) = g\Big\{1,\frac{1}{N}(k+\tfrac{1}{2})z\Big\} = \exp\Big[(k+\tfrac{1}{2})\Big\{z-\sum_{p=1}^{\infty}\frac{c_{2p}}{2p}(\alpha z)^{2p}\Big\}\Big]$$

$$= \sum_{\nu=0}^{\infty}\frac{1}{\nu!}A_\nu(\alpha,k)z^\nu, \quad (4.7)$$

where

$$C_{2p} = (-1)^{p-1}\frac{B_p}{(2p+1)!}\Big[\sum_{\mu=0}^{m-1}2^{-2\mu}D_{2\mu}\binom{2p+1}{2\mu}(k+\tfrac{1}{2})^{-2\mu}+$$

$$+B_m\delta_p\binom{2p+1}{2m}\Big(1+\frac{1}{2k+1}\Big)^{2p-2m+1}(k+\tfrac{1}{2})^{-2m}\Big], \quad (4.8)$$

and $$A_\nu(\alpha,k) = b_\nu(N,k)\Big(\frac{k+\frac{1}{2}}{N}\Big)^\nu = b_\nu(N,k)\Big(\frac{\alpha}{k+\frac{1}{2}}\Big)^\nu. \quad (4.9)$$

We apply Theorem 3 to $G(1,z)$, writing in (3.6)

$$\sigma = 1, \qquad \tau = \tfrac{1}{2}, \qquad h = k,$$

$$C_{2p} = (-1)^{p-1}\frac{B_p}{(2p+1)!}, \qquad C_{2p-1} = 0, \qquad E_\mu = 2^{-2\mu}D_{2\mu}, \qquad a = 1.$$

We have

$$t = u\Big(1+\sum_{p=1}^{\infty}(-1)^{p-1}\frac{B_p}{(2p+1)!}u^{2p}\Big)^{-1} = \Big(\tfrac{1}{4}+u^{-2}\int_0^u \frac{x}{e^x-1}\,dx\Big)^{-1},$$

$$\int_0^{\alpha}\{u(t)-t\}t^{-2}\,dt = 2-2\frac{u(\alpha)}{\alpha}-\log\alpha+\log\{2\sinh\tfrac{1}{2}u(\alpha)\},$$

by (1.7). Hence Theorem 3 gives

$$\log A_{k-1}(\alpha,k) = (k-1)\log(k+\tfrac{1}{2})-$$
$$-(k+\tfrac{1}{2})\Big(2-2\frac{u(\alpha)}{\alpha}-\log\alpha+\log\{2\sinh\tfrac{1}{2}u(\alpha)\}\Big)+$$
$$+\tfrac{1}{2}\log\frac{u(\alpha)}{\alpha}+\tfrac{1}{2}\log u'(\alpha)+\sum_{\mu=1}^{m-1}\phi_\mu(\alpha)k^{-\mu}+O(k^{-m}): \quad (4.10)$$

that is, by (4.9),

$$\log b_{k-1}(N,k) = (k-1)\log(k+\tfrac{1}{2})-(k-1)\log\alpha+\log A_{k-1}(\alpha,k)$$
$$= 2(k-1)\log(k+\tfrac{1}{2})+(k+\tfrac{1}{2})\Big(2\frac{u(\alpha)}{\alpha}-\log\{2\sinh\tfrac{1}{2}u(\alpha)\}\Big)-$$
$$-2(k+\tfrac{1}{2})+\log\alpha+\tfrac{1}{2}\log\{uu'(\alpha)\}+\sum_{\mu=1}^{m-1}\phi_\mu(\alpha)k^{-\mu}+O(k^{-m}).$$

Stirling's formula implies

$$\log\{k!(k-1)!\}$$
$$= 2k\log k-2k+\log(2\pi)+\sum_{\mu=1}^{m-1}(-1)^{\mu-1}\frac{B_\mu}{\mu(2\mu-1)}k^{-2\mu+1}+O(k^{-2m+1}).$$

Hence, by (4.3),

$$\log R_1(N,k) = \log b_{k-1}(N,k)-\log\{k!(k-1)!\}$$
$$= (k+\tfrac{1}{2})\Big[2\frac{u(\alpha)}{\alpha}-\log\{2\sinh\tfrac{1}{2}u(\alpha)\}\Big]+2k\log\Big(1+\frac{1}{2k}\Big)-1-2\log(k+\tfrac{1}{2})+$$
$$+\log\alpha-\log(2\pi)+\tfrac{1}{2}\log\{uu'(\alpha)\}+$$
$$+\sum_{\nu=1}^{[m/2]}(-1)^{\nu}\frac{B_\nu}{\nu(2\nu-1)}k^{-2\nu+1}+\sum_{\mu=1}^{m-1}\phi_\mu(\alpha)k^{-\mu}+O(k^{-m})$$
$$= k\Big[2\frac{u(\alpha)}{\alpha}-\log\{2\sinh\tfrac{1}{2}u(\alpha)\}\Big]-\log N+\frac{u(\alpha)}{\alpha}-\tfrac{1}{2}\log\Big(\frac{2}{u(\alpha)}\sinh\tfrac{1}{2}u(\alpha)\Big)+$$
$$+\tfrac{1}{2}\log u'(\alpha)-\log(2\pi)+\sum_{\mu=1}^{m-1}k^{-\mu}\psi_\mu(\alpha)+O(k^{-m}).$$

This proves the statement. The range of validity of the expression is given by the radius of convergence of $v(t)$ defined by

$$t = v\Big(1+\sum_{p=1}^{\infty}\frac{B_p}{(2p+1)!}v^{2p}\Big)^{-1} = v^2\Big(\int_0^v(2-\tfrac{1}{2}z\cot\tfrac{1}{2}z)\,dz\Big)^{-1}.$$

Since the coefficients $B_p/(2p+1)!$ are positive,

$$\left|\int_0^v (2-\tfrac{1}{2}z\cot\tfrac{1}{2}z)\,dz\right| \leqslant \int_0^r (2-\tfrac{1}{2}x\cot\tfrac{1}{2}x)\,dx$$

on the circle $|v| = r$. Hence the minimum of $|t(v)|$ on $|v| = r$ is

$$M(r) = r^2\left(\int_0^r (2-\tfrac{1}{2}\cot\tfrac{1}{2}x)\,dx\right)^{-1}.$$

It follows from the Bürmann–Lagrange theorem* that $v(t)$ is certainly convergent in a circle with radius $M(r)$. The maximum $\rho = M(R)$ of $M(r)$ is where

$$\tfrac{1}{2}R(2-\tfrac{1}{2}R\cot\tfrac{1}{2}R) = \int_0^R (2-\tfrac{1}{2}x\cot\tfrac{1}{2}x)\,dx,$$

i.e. at $R = 4{\cdot}2048\ldots$. Hence

$$\rho = R^2\left(\int_0^R (2-\tfrac{1}{2}x\cot\tfrac{1}{2}x)\,dx\right)^{-1} = R(1-\tfrac{1}{4}R\cot\tfrac{1}{2}R)^{-1} = 2{\cdot}5984\ldots.$$

5. The circulatory part

I have to show now that the circulatory part has a smaller order of magnitude than the principal term $R_1(N,k)$. This corresponds to the fact (which, however, I shall not use) that $z = 1$ is the 'strongest' singularity of the generating function

$$\{(1-z)(1-z^2)\ldots(1-z^k)\}^{-1} = 1+\sum_{n=1}^{\infty} P(n,k)z^n.$$

In fact, I shall prove that

$$\sum_{q=2}^{k}\sum_{\zeta} R_\zeta(N,k) = o\{k^{-m}R_1(N,k)\}$$

for every $m > 0$.

We first consider the dominating term $R_{-1}(N,k)$, which can be treated very similarly to $R_1(N,k)$. We have from (3.8)

$$f(-1,t) = (-1)^{-N+\frac{1}{2}k(k+1)}2^{-k}(h!)^{-1}e^{Nt}\prod_{\nu=1}^{h'}\{\cosh(\nu-\tfrac{1}{2})t\}^{-1}\prod_{\nu=1}^{h}\nu t(\sinh\nu t)^{-1},$$

$$h = [\tfrac{1}{2}k], \qquad h' = [\tfrac{1}{2}k+\tfrac{1}{2}],$$

and $$R_{-1}(N,k) = (-1)^{-N+\frac{1}{2}k(k+1)}2^{-k}\{h!(h-1)!\}^{-1}b_{h-1}(-1,N,k), \qquad (5.1)$$

* See Hurwitz–Courant, *Funktionentheorie*, 2. Aufl. (Berlin, 1925), 139.

where

$$g(-1,t) \equiv \sum_{\nu=0}^{\infty} \frac{1}{\nu!} b_\nu(-1,N,k)t^\nu$$

$$= \exp\left\{Nt - \sum_{\nu=1}^{h'} \log\{\cosh(\nu-\tfrac{1}{2})t\} - \sum_{\nu=1}^{h} \log\left(\frac{\sinh \nu t}{\nu t}\right)\right\}$$

$$= \exp\left\{Nt - \sum_{\nu=1}^{h'} \sum_{p=1}^{\infty} (-1)^{p-1} \frac{(2^{2p}-1)B_p}{2p(2p)!}[(2\nu-1)t]^{2p} - \right.$$

$$\left. - \sum_{\nu=1}^{h} \sum_{p=1}^{\infty} (-1)^{p-1} \frac{B_p}{2p(2p)!}(2\nu t)^{2p}\right\}$$

$$= \exp\left\{Nt - \sum_{p=1}^{\infty} \left[\sum_{\nu=1}^{h'} (2\nu-1)^{2p}\right](-1)^{p-1} \frac{(2^{2p}-1)B_p}{2p(2p)!} t^{2p} - \right.$$

$$\left. - \sum_{p=1}^{\infty} \left[\sum_{\nu=1}^{h} (2\nu)^{2p}\right](-1)^{p-1} \frac{B_p}{2p(2p)!} t^{2p}\right\}$$

$$= \exp\left\{Nt - \tfrac{1}{2}(k+\tfrac{1}{2}) \sum_{p=1}^{\infty} (-1)^{p-1} \frac{B_p}{2p(2p+1)!} \times \right.$$

$$\left. \times [1+\delta_p(2p+1)k^{-1}][2(k+\tfrac{1}{2})t]^{2p}\right\} \quad (|\delta_p| < 1).$$

The Euler formula is used here in the form

$$\sum_{\nu=1}^{[\frac{1}{2}k+\frac{1}{4}]} (\nu-\tfrac{1}{4})^{2p}, \qquad \sum_{\nu=1}^{[\frac{1}{2}k]} \nu^{2p} = \frac{1}{2p+1}\{\tfrac{1}{2}(k+\tfrac{1}{2})\}^{2p+1} + \delta\{\tfrac{1}{2}(k+\tfrac{1}{2})\}^{2p} \quad (|\delta| < 1).$$

We change the variable t into

$$z = \frac{2N}{k+\frac{1}{2}} t = \frac{2k+1}{\alpha} t,$$

and apply the corollary of Theorem 3 to

$$\exp\left\{\tfrac{1}{2}(k+\tfrac{1}{2})\left(z - \sum_{p=1}^{\infty} (-1)^{p-1} \frac{B_p}{2p(2p+1)!}\{1+\delta_p(2p+1)k^{-1}\}(\alpha z)^{2p}\right)\right\}.$$

We obtain as in the previous section that

$$\log\left\{b_{h-1}(-1,N,k)\left(\frac{\alpha}{2k+1}\right)^{h-1}\right\} = (h-1)\log\{\tfrac{1}{2}(k+\tfrac{1}{2})\} - \tfrac{1}{2}(k+\tfrac{1}{2}) \times$$

$$\times\left\{2 - 2\frac{u(\alpha)}{\alpha} + \log\left(\frac{2}{\alpha}\sinh \tfrac{1}{2}u(\alpha)\right)\right\} + O(1),$$

$$\log b_{h-1}(-1,N,k) = 2(h-1)\log(k+\tfrac{1}{2}) - (h-1)\log\alpha -$$

$$-h\left\{2 - 2\frac{u(\alpha)}{\alpha} + \log\left(\frac{2}{a}\sinh \tfrac{1}{2}u(\alpha)\right)\right\} + O(1).$$

Hence, by (5.1) and Stirling's formula,

$$\log\{(-1)^{N-\frac{1}{2}k(k+1)}R_{-1}(N,k)\}$$
$$= -k\log 2-\log\{h!(h-1)!\}+\log b_{h-1}(-1,N,k)$$
$$= h\left\{2\frac{u(\alpha)}{\alpha}-\log\left(2\sinh\tfrac{1}{2}u(\alpha)\right)\right\}-\log N+O(1).$$

Since

$$2\frac{u(\alpha)}{\alpha}-\log\{2\sinh\tfrac{1}{2}u(\alpha)\} > 2\frac{u(\alpha)}{\alpha}-\tfrac{1}{2}u(\alpha)$$
$$= 2u(\alpha)\left(\frac{1}{\alpha}-\frac{1}{u}\right) = \frac{2}{u}\int_0^u \frac{x}{e^x-1}\,dx > 0$$

and $h \cong \frac{1}{2}k$, it follows that $R_{-1}(N,k)$ has a smaller exponential order than $R_1(N,k)$ and that $R_{-1}(N,k) = o\{k^{-m}R_1(N,k)\}$ for every $m > 0$.

6. The circulatory part (continued)

Suppose now that $q > 2$ and write

$$h = \left[\frac{k}{q}\right], \qquad k = qh+r \quad (0 \leqslant r < q),$$

$$\pi(r,\zeta) = (1-\zeta)(1-\zeta^2)\ldots(1-\zeta^r) \quad \text{if} \quad 0 < r < q,\ \pi(0,\zeta) = 1. \quad (6.1)$$

Then

$$\pi(q-1,\zeta) = (1-\zeta)\ldots(1-\zeta^{q-1}) = \lim_{x\to 1}\frac{x^q-1}{x-1} = q,$$

and, from (3.8),

$$f(\zeta,t) = \zeta^{-N+\frac{1}{2}k(k+1)}\{\pi(r,\zeta)q^{2h}h!\}^{-1}t^{-h}e^{Nt}\times$$
$$\times \prod_{j=1}^{k}{}' (1-\zeta^j)(e^{\frac{1}{2}jt}-\zeta^j e^{-\frac{1}{2}jt})^{-1}\prod_{\nu=1}^{h}\tfrac{1}{2}qvt(\sinh\tfrac{1}{2}qvt)^{-1}.$$

Here Π' signifies that the product index j takes only those values which are not multiples of q. A similar remark applies to Σ' below.

Write

$$g(\zeta,t) = \sum_{\nu=0}^{\infty}\frac{1}{\nu!}b_\nu(\zeta,N,k)t^\nu = \exp\Big\{Nt-\sum_{j=1}^{k}{}'\log[(1-\zeta^j)^{-1}(e^{\frac{1}{2}jt}-\zeta^j e^{-\frac{1}{2}jt})]-$$
$$-\sum_{\nu=1}^{h}\log[(\tfrac{1}{2}qvt)^{-1}\sinh\tfrac{1}{2}qvt]\Big\}. \quad (6.2)$$

Then

$$R_\zeta(N,k) = \zeta^{-N+\frac{1}{2}k(k+1)}\{\pi(r,\zeta)q^{2h}h!(h-1)!\}^{-1}b_{h-1}(\zeta,N,k).$$

Now, if $l = [\frac{1}{6}q]$,

$$|\pi(r,\zeta)| = |(1-\zeta)\ldots(1-\zeta^r)| \geqslant |(1-e^{2\pi i/q})(1-e^{4\pi i/q})\ldots(1-e^{2l\pi i/q})|$$
$$\geqslant \prod_{\lambda=1}^{l}\frac{6\lambda}{q} = \left(\frac{6}{q}\right)^l l! \geqslant e^{-\frac{1}{6}q}.$$

Hence $$|R_\zeta(N,k)| \leqslant q^{-2h}\{h!(h-1)!\}^{-1}e^{\frac{1}{2}q}|b_{h-1}(\zeta,N,k)|. \tag{6.3}$$

To obtain an estimate for b_{h-1} we have to study the function in the exponent of (6.2). Let us write, for $\beta \neq 1$,

$$\phi(\beta,t) = -\log\frac{e^{\frac{1}{2}t}-\beta e^{-\frac{1}{2}t}}{1-\beta} = \sum_{\mu=1}^{\infty}\frac{1}{\mu!}B(\beta,\mu)t^\mu. \tag{6.4}$$

We have

$$B(\beta,1) = \frac{\beta+1}{2(\beta-1)},$$

$$B(\beta,\mu+2) = (\beta-1)^{-\mu-2}(L_{\mu 0}+L_{\mu 1}\beta+\ldots+L_{\mu\mu}\beta^\mu)\beta \quad (\mu \geqslant 0), \tag{6.5}$$

where the numbers $L_{\mu\nu}$ are positive integers satisfying the recursion

$$L_{\mu+1,\nu+1} = (\mu+1-\nu)L_{\mu\nu}+(\nu+2)L_{\mu,\nu+1} \quad (\mu > \nu \geqslant 0),$$

$$L_{00} = L_{\mu 0} = L_{\mu\mu} = 1. \tag{6.6}$$

For $$\phi_t'(\beta,t) = \frac{(\beta-e^t)}{2(\beta+e^t)}$$

and $$\phi_t^{(\mu+2)}(\beta,t) = (\beta-e^t)^{-\mu-2}\beta e^t\sum_{\nu=0}^{\mu}L_{\mu\nu}\beta^\nu e^{(\mu-\nu)t} \quad (\mu \geqslant 0),$$

as is easily verified by induction.

Writing $t = (\beta-1)x$ and making $\beta \to 1$ in (6.4) and (6.5), we obtain, since

$$\lim_{\beta\to 1}(1-\beta)^{-1}\{e^{\frac{1}{2}x(\beta-1)}-\beta e^{-\frac{1}{2}x(\beta-1)}\} = 1-x,$$

$$\frac{1}{(\mu+2)!}(L_{\mu 0}+L_{\mu 1}+\ldots+L_{\mu\mu}) = \frac{1}{\mu+2},$$

i.e. $$L_{\mu 0}+L_{\mu 1}+\ldots+L_{\mu\mu} = (\mu+1)!.$$

Hence for an arbitrary primitive qth root of unity ζ $(q \geqslant 3)$

$$|B(\zeta^j,\mu)| \leqslant |1-\zeta^j|^{-\mu}(\mu-1)! \leqslant \left(\frac{q}{3\sqrt{3}}\right)^\mu(\mu-1)!. \tag{6.7}$$

Next I show that

$$\sum_{j=1}^{q-1}B(\zeta^j,2\nu-1) = 0, \qquad \sum_{j=1}^{q-1}B(\zeta^j,2\nu) = (-1)^\nu\frac{B_\nu}{2\nu}(q^{2\nu}-1) \quad (\nu = 1,2,\ldots). \tag{6.8}$$

Using the identity

$$\prod_{j=1}^{q-1}(a-\zeta^j) = \frac{a^q-1}{a-1} \quad (a \neq 1),$$

we obtain

$$\sum_{j=1}^{q-1} \log \frac{e^{\frac{1}{2}t}-\zeta^j e^{-\frac{1}{2}t}}{1-\zeta^j} = -\tfrac{1}{2}(q-1)t + \sum_{j=1}^{q-1} \log \frac{e^t-\zeta^j}{1-\zeta^j}$$

$$= -\tfrac{1}{2}(q-1)t + \log \prod_{j=1}^{q-1} \frac{e^t-\zeta^j}{1-\zeta^j} = -\tfrac{1}{2}(q-1)t + \log \frac{e^{qt}-1}{q(e^t-1)}$$

$$= \log\left(\frac{2}{qt}\sinh \tfrac{1}{2}qt\right) - \log\left(\frac{2}{t}\sinh \tfrac{1}{2}t\right) = \sum_{\nu=1}^{\infty} (-1)^{\nu-1} \frac{B_\nu}{2\nu(2\nu)!}(q^{2\nu}-1)t^{2\nu}.$$

Then (6.8) follows from this and (6.4).

Finally we note the following formula:

Let $0 \leqslant j < q$, $h' = [(k-j)/q]$, *then*

$$\sum_{\nu=0}^{h'} (q\nu+j)^p = \frac{1}{q(p+1)}(k+\tfrac{1}{2})^{p+1} + \delta(k+q)^p \quad (|\delta| < 1). \qquad (6.9)$$

This follows from Euler's formula or simply by comparing

$$\sum_{\nu=0}^{h'} \left(\nu+\frac{j}{q}\right)^p \quad \text{with} \quad \int_0^{(1/q)(k+\frac{1}{2})} x^p\,dx.$$

Now we can develop the function in the exponent of (6.2) into a power series. For the first sum we have

$$-\sum_{j=1}^{k}{}' \log\{(1-\zeta^j)^{-1}(e^{\frac{1}{2}jt}-\zeta^j e^{-\frac{1}{2}jt})\}$$

$$= \sum_{j=1}^{k}{}' \phi(\zeta^j, jt) = \sum_{j=1}^{k}{}' \sum_{p=1}^{\infty} \frac{1}{p!} B(\zeta^j, p) j^p t^p, \quad \text{by (6.4)},$$

$$= \sum_{p=1}^{\infty} \frac{1}{p!} t^p \sum_{j=1}^{q-1} B(\zeta^j, p) \sum_{\nu=0}^{h'} (q\nu+j)^p$$

$$= \sum_{p=1}^{\infty} \frac{1}{p!} t^p \sum_{j=1}^{q-1} B(\zeta^j, p)\left\{\frac{1}{q(p+1)}(k+\tfrac{1}{2})^{p+1} + \delta(k+q)^p\right\}, \quad \text{by (6.9)},$$

$$= \frac{k+\frac{1}{2}}{q} \sum_{\nu=1}^{\infty} (-1)^\nu \frac{B_\nu}{2\nu(2\nu+1)!}(q^{2\nu}-1)\{(k+\tfrac{1}{2})t\}^{2\nu} +$$

$$+ (q-1) \sum_{p=1}^{\infty} \frac{\delta_p}{p}\left\{\frac{q}{3\sqrt{3}}(k+q)t\right\}^p, \quad \text{by (6.7), (6.8)}.$$

For the second sum we have

$$-\sum_{\nu=1}^{h}\log\{(\tfrac{1}{2}q\nu t)^{-1}\sinh\tfrac{1}{2}q\nu t\}$$

$$=\sum_{\nu=1}^{h}\sum_{p=1}^{\infty}(-1)^p\frac{B_p}{2p(2p)!}(q\nu t)^{2p}=\sum_{p=1}^{\infty}(-1)^p\frac{B_p}{2p(2p)!}t^{2p}\sum_{\nu=1}^{h}(q\nu)^{2p}$$

$$=\frac{k+\frac{1}{2}}{q}\sum_{p=1}^{\infty}(-1)^p\frac{B_p}{2p(2p+1)!}\{(k+\tfrac{1}{2})t\}^{2p}+\sum_{p=1}^{\infty}\delta_p\frac{B_p}{2p(2p)!}t^{2p}(k+q)^{2p}$$

$$=\frac{k+\frac{1}{2}}{q}\sum_{p=1}^{\infty}(-1)^p\frac{B_p}{2p(2p+1)!}\{(k+\tfrac{1}{2})t\}^{2p}+\sum_{p=1}^{\infty}\frac{\delta_p}{p}\left\{\frac{q}{3\sqrt{3}}(k+q)t\right\}^p$$

since $$B_p/(2p)!<(q/3\sqrt{3})^{2p}.$$

Combining these results with (6.2), we have

$$g(\zeta,t)=\sum_{\nu=0}^{\infty}\frac{1}{\nu!}b_\nu(\zeta,N,k)t^\nu$$

$$=\exp\left\{Nt-\frac{k+\frac{1}{2}}{q}\sum_{p=1}^{\infty}(-1)^{p-1}\frac{B_p}{2p(2p+1)!}[q(k+\tfrac{1}{2})t]^{2p}+\right.$$

$$\left.+q\sum_{p=1}^{\infty}\frac{\delta_p}{p}\left(\frac{q}{3\sqrt{3}}(k+q)t\right)^p\right\}.$$

Changing t into

$$z=\frac{Nq}{k+\frac{1}{2}}t=\frac{k+\frac{1}{2}}{\alpha}qt$$

and noting that

$$\alpha<\rho=2{\cdot}5984\ldots<\frac{3\sqrt{3}(k+\tfrac{1}{2})}{k+q},^*\qquad\frac{\alpha(k+q)}{3\sqrt{3}(k+\frac{1}{2})}<1,$$

we obtain

$$\sum_{\nu=0}^{\infty}\frac{1}{\nu!}\left\{\frac{\alpha}{q(k+\frac{1}{2})}\right\}^\nu b_\nu(\zeta,N,k)z^\nu=G(q,z)\exp\left\{q\sum_{p=1}^{\infty}\frac{\delta_p}{p}\left(\frac{\alpha(k+q)}{3\sqrt{3}(k+\frac{1}{2})}z\right)^p\right\}$$

$$=G(q,z)\exp\left\{q\sum_{p=1}^{\infty}\frac{\delta_p}{p}z^p\right\}$$

$$=G(q,z)\sum_{\mu=0}^{\infty}\delta_\mu\binom{q+\mu-1}{\mu}z^\mu,$$

* As a matter of fact, $\rho=2{\cdot}5984\ldots$ is slightly bigger than $\frac{1}{2}3\sqrt{3}$ and the above inequality is not true if $h=1$ and q is very near to k. But, if $q\geqslant 4$, then obviously $3\sqrt{3}$ can be replaced by $4\sqrt{2}$ in (6.7), and the inequality becomes $\alpha<2\sqrt{2}=2{\cdot}8\ldots$.

where

$$G(q,z)=\exp\left\{\frac{1}{q}(k+\tfrac{1}{2})\left(z-\sum_{p=1}^{\infty}(-1)^{p-1}\frac{B_p}{2p(2p+1)!}(\alpha z)^{2p}\right)\right\}=\sum_{\nu=0}^{\infty}\frac{1}{\nu!}A_\nu z^\nu. \tag{6.10}$$

Hence

$$\frac{1}{(h-1)!}|b_{h-1}(\zeta,N,k)|\leqslant\left\{\frac{q}{\alpha}(k+\tfrac{1}{2})\right\}^{h-1}\sum_{\mu+\nu=h-1}\binom{q+\mu-1}{\mu}\frac{1}{\nu!}A_\nu. \tag{6.11}$$

Let us put $K=(k+\frac{1}{2})/q$ and apply Theorem 2 to the coefficients

$$A_{l-1}, \qquad l=\sigma K \quad (\sigma\leqslant 1).$$

We have, for $u=u(\sigma,t)$,

$$t=u\left\{\sigma+\sum_{p=1}^{\infty}(-1)^{p-1}\frac{B_p}{(2p+1)!}u^{2p}\right\}^{-1}=u\left(\sigma-1+\frac{1}{u}\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx\right)^{-1} \tag{6.12}$$

and, from (3.5),

$$\begin{aligned}\log A_{l-1}&=(l-1)\log K-K\int_0^\alpha\frac{u(\sigma,t)-\sigma t}{t^2}\,dt+O(1)\\&=(\sigma K-1)\log K-K\left\{2\sigma+\sigma\log\frac{u(\sigma,\alpha)}{\sigma\alpha}-2\frac{u(\sigma,\alpha)}{\alpha}+\right.\\&\qquad\left.+\log\left(\frac{2}{u(\sigma,\alpha)}\sinh\tfrac{1}{2}u(\sigma,\alpha)\right)\right\}+O(1),\end{aligned} \tag{6.13}$$

since, by (6.12), $\dfrac{\partial u}{\partial\alpha}\left(\dfrac{2}{\alpha}+\dfrac{1-\sigma}{u}-\tfrac{1}{2}\coth\tfrac{1}{2}u\right)=u\alpha^{-2}$.

Stirling's formula gives

$$\begin{aligned}\log\left\{\frac{1}{(l-1)!}A_{l-1}\right\}&=\log A_{l-1}-(l-\tfrac{1}{2})\log l+l+O(1)\\&=K\left\{2\frac{u(\sigma,\alpha)}{\alpha}-\sigma-\sigma\log\frac{u(\sigma,\alpha)}{\alpha}-\log\left[\frac{2}{u(\sigma,\alpha)}\sinh\tfrac{1}{2}u(\sigma,\alpha)\right]\right\}-\\&\qquad-\log K+\tfrac{1}{2}\log l+O(1).\end{aligned} \tag{6.14}$$

The expression in the brackets $\{\}$ has a maximum (at fixed α) if

$$\frac{\partial}{\partial\sigma}\left\{2\frac{u(\sigma,\alpha)}{\alpha}-\sigma-\sigma\log\frac{u(\sigma,\alpha)}{\alpha}-\log\left(\frac{2}{u(\sigma,\alpha)}\sinh\tfrac{1}{2}u(\sigma,\alpha)\right)\right\}=0,$$

$$\frac{\partial u}{\partial\sigma}\left\{\frac{2}{\alpha}+\frac{1-\sigma}{u}-\tfrac{1}{2}\coth\tfrac{1}{2}u\right\}=1+\log\frac{u}{\alpha}.$$

Here the left-hand side is 1, as is easily seen from (6.12). Hence (6.14) is maximum if $u(\sigma,\alpha)=\alpha$, i.e., by (6.12), if

$$\sigma = 1-\sum_{p=1}^{\infty}(-1)^{p-1}\frac{B_p}{(2p+1)!}\alpha^{2p} = 2-\frac{1}{\alpha}\int_0^{\alpha}\tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx.$$

For this value of σ, the expression in (6.14) becomes

$$2-\sigma-\log\left(\frac{2}{u}\sinh\tfrac{1}{2}u\right) = \frac{1}{\alpha}\int_0^{\alpha}\tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx-\log\left(\frac{2}{\alpha}\sinh\tfrac{1}{2}\alpha\right),$$

and

$$\log\left\{\frac{1}{(l-1)!}A_{l-1}\right\} < K\left\{\frac{1}{\alpha}\int_0^{\alpha}\tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx-\log\left(\frac{2}{\alpha}\sinh\tfrac{1}{2}\alpha\right)\right\}.$$

Combining this with (6.11) and (6.3), and noting that

$$K=\frac{1}{q}(k+\tfrac{1}{2}) < h-1$$

and

$$\binom{q+\mu-1}{\mu} < \binom{q+h}{h} = \frac{(q+h)!}{q!\,h!},$$

in (6.11), we obtain

$$\begin{aligned}\log|R_\zeta(N,k)| &< -2h\log q-(h+\tfrac{1}{2})\log h+h+\tfrac{1}{6}q+(h-1)\log\left\{\frac{q}{\alpha}(k+\tfrac{1}{2})\right\}+\\ &\quad+(q+h)\log(q+h)-q\log q-h\log h+\\ &\quad+h\left\{\frac{1}{\alpha}\int_0^{\alpha}\tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx-\log\left(\frac{2}{\alpha}\sinh\tfrac{1}{2}\alpha\right)\right\}\\ &= h\left\{1+\frac{1}{\alpha}\int_0^{\alpha}\tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx-\log(2\sinh\tfrac{1}{2}\alpha)\right\}-\log N+h\log(1+q/h)+\\ &\quad+q\log(1+h/q)+\tfrac{1}{6}q+O(\log h).\end{aligned}$$

Here $h\log(1+q/h)=O(k^{\frac{1}{2}}), \qquad q\log(1+h/q)=O(k^{\frac{1}{2}})$

since $\left[\frac{k}{gh}\right]=1.$

Hence

$$\begin{aligned}\log|R_\zeta(N,k)| &< h\left\{1+\frac{1}{\alpha}\int_0^{\alpha}\tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx-\log(2\sinh\tfrac{1}{2}\alpha)\right\}-\\ &\quad-\log N+\tfrac{1}{6}q+O(k^{\frac{1}{2}})\\ &\leqslant \tfrac{1}{2}k\left\{1+\frac{1}{\alpha}\int_0^{\alpha}\tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx-\log(2\sinh\tfrac{1}{2}\alpha)+\tfrac{1}{2}\right\}-\\ &\quad-\log N+O(k^{\frac{1}{2}}).\end{aligned}$$

In § 4 we have proved that

$$\log R_1(N,k) = k\left\{2\frac{u(\alpha)}{\alpha} - \log[2\sinh \tfrac{1}{2}u(\alpha)]\right\} - \log N + O(1).$$

Hence we have finished the proof if we can show that

$$\frac{3}{2} + \frac{1}{\alpha}\int_0^\alpha \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx - \log(2\sinh\tfrac{1}{2}\alpha) < 3\left\{2\frac{u(\alpha)}{\alpha} - \log[2\sinh\tfrac{1}{2}u(\alpha)]\right\}$$

$$= 3\left\{\frac{2}{u}\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx - \log[2\sinh\tfrac{1}{2}u(\alpha)]\right\}. \quad (6.15)$$

Now, if $u > 0$, then

(i) $\dfrac{1}{u}\displaystyle\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx$ is monotone increasing,

(ii) $\dfrac{1}{u}\displaystyle\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx - \log\left(\frac{2}{u}\sinh\tfrac{1}{2}u\right)$ is monotone decreasing,

(iii) $\dfrac{2}{u}\displaystyle\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx - \log\left(\frac{2}{u}\sinh\tfrac{1}{2}u\right)$ is monotone increasing with u.

For,

(i) $$\frac{d}{du}\left(\frac{1}{u}\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx\right) = \frac{1}{u}\left(\tfrac{1}{2}u\coth\tfrac{1}{2}u - \frac{1}{u}\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx\right) > 0$$

if $u > 0$,

(ii) $$\frac{d}{du}\left\{\frac{1}{u}\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx - \log\left(\frac{2}{u}\sinh\tfrac{1}{2}u\right)\right\}$$

$$= \frac{1}{u}\left(1 - \frac{1}{u}\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx\right) < 0 \quad \text{if } u > 0,$$

(iii) $$\frac{d}{du}\left\{\frac{2}{u}\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx - \log\left(\frac{2}{u}\sinh\tfrac{1}{2}u\right)\right\}$$

$$= \frac{1}{u^2}\left(\tfrac{1}{2}u^2\coth\tfrac{1}{2}u + u - 2\int_0^u \tfrac{1}{2}x\coth\tfrac{1}{2}x\,dx\right)$$

and the last expression is positive for $u > 0$, since

$$\frac{d}{du}\left(\tfrac{1}{2}u^2 \coth \tfrac{1}{2}u + u - 2\int_0^u \tfrac{1}{2}x \coth \tfrac{1}{2}x\, dx\right) = 1 - \tfrac{1}{4}u^2(\operatorname{cosech} \tfrac{1}{2}u)^2 > 0.$$

Hence

$$\frac{3}{2} + \frac{1}{\alpha}\int_0^\alpha \tfrac{1}{2}x \coth \tfrac{1}{2}x\, dx - \log\left(\frac{2}{\alpha}\sinh \tfrac{1}{2}\alpha\right) \leqslant \frac{5}{2} \quad \text{by (ii)},$$

$$3\left\{\frac{2}{u}\int_0^u \tfrac{1}{2}x \coth \tfrac{1}{2}x\, dx - \log\left(\frac{2}{u}\sinh \tfrac{1}{2}u\right)\right\} \geqslant 6 \quad \text{by (iii)},$$

and (6.15) is true if

$$\tfrac{5}{2} - \log\alpha < 6 - 3\log u(\alpha), \qquad 3\log u(\alpha) - \log\alpha < \tfrac{7}{2}.$$

Now

$$\frac{u(\alpha)}{\alpha} = \frac{1}{u}\int_0^u \tfrac{1}{2}x \coth \tfrac{1}{2}x\, dx$$

is increasing with u (and α) by (i). Hence $3\log u(\alpha) - \log\alpha$ is monotone increasing with α. But, if $u(\alpha) = 3{\cdot}5$, then

$$\alpha = 3{\cdot}5^2\left(\int_0^{3\cdot5} \tfrac{1}{2}x \coth \tfrac{1}{2}x\, dx\right)^{-1} = 2{\cdot}68\ldots > \rho,$$

$$3\log u(\alpha) - \log\alpha < 3\log 3{\cdot}5 - \log 2{\cdot}68 < \tfrac{7}{2},$$

which proves (6.15).

[*Note added 14 July 1950.*]

From this point of view it is interesting to compare by a numerical example the Sylvester terms for $P(n, n)$ and the Hardy–Ramanujan terms for $P(n)$. For $q \leqslant n$, write $P_q(n) = \sum_\zeta P_\zeta(n, n)$, where ζ runs through the primitive qth roots of unity. Taking $n = 8$ we obtain

$$P_1(8) = 21{\cdot}4127, \quad P_2(8) = 0{\cdot}4112, \quad P_3(8) = -0{\cdot}0566, \quad P_4(8) = 0{\cdot}1016,$$
$$P_5(8) = 0, \quad P_6(8) = 0{\cdot}0278, \quad P_7(8) = 0{\cdot}0408, \quad P_8(8) = 0{\cdot}0625,$$

whereas the first five terms of the modified exact Hardy–Ramanujan–Rademacher series are

$$21{\cdot}7092 + 0{\cdot}3463 - 0{\cdot}0896 + 0{\cdot}0500 - 0{\cdot}0192 + \ldots.$$

Of course, the sum of both series is $P(8) = 22$.

For large values of n, I suspect a very close approximation of the $P_q(n)$ by the corresponding Hardy–Ramanujan–Rademacher terms. Unfortunately, it is not easy to obtain data for comparison since the computation of the Sylvester waves becomes almost prohibitive for large values of n. In one instance I have calculated $P_2(40)$ and found $+7{\cdot}4784$, whereas the second Rademacher term for $P(40)$ is $+7{\cdot}4311$. In the computation I was kindly assisted by the Mathematical Statistics section of the Commonwealth Scientific and Industrial Research Organization.

OVERCONVERGENCE PROPERTIES OF SOME INTERPOLATION SERIES

By SHEILA SCOTT MACINTYRE (*Aberdeen*)

[Received 3 February 1950]

I. Lidstone Series

1. Classical results connected with the Lidstone expansion

$$\sum_{i=0}^{\infty}[f^{(2i)}(a)\alpha_i(z)+f^{(2i)}(b)\beta_i(z)] \tag{1}$$

are embodied in Theorem I [(**6**) 287 and (**8**) 455] and Theorem II [(**6**) 292 and (**8**) 455] below. $M(r)$ is the maximum modulus of $f(z)$ on $|z| = r$. By a *derivative of order* 0 is understood, as usual, the function itself, and $\alpha_n(z)$, $\beta_n(z)$ are the Lidstone polynomials defined by the formulae

$$\alpha_n''(z) = \alpha_{n-1}(z), \qquad \beta_n''(z) = \beta_{n-1}(z) \quad (n \geqslant 1) \tag{2}$$

together with

$$\alpha_0(z) = (z-b)/(a-b), \qquad \beta_0(z) = (z-a)/(b-a), \tag{3}$$

$$\alpha_n(a) = \alpha_n(b) = \beta_n(a) = \beta_n(b) = 0 \quad (n \geqslant 1). \tag{4}$$

Alternative definitions due to Poritsky and equivalent to (2), (3), (4) are contained in (7), (8), (9).

Theorem I. *If $f(z)$ is an integral function satisfying*

$$f(z) = O(e^{k|z|}) \quad (0 < k < \pi/|a-b|), \tag{5}$$

then the Lidstone series (1) *converges to $f(z)$ for all z, the convergence being uniform in any bounded region.*

Theorem II. *If the series*

$$\sum_{i=0}^{\infty}[a_i\alpha_i(z)+b_i\beta_i(z)]$$

converges for any value of z, not congruent to a $(\mathrm{mod}(b-a))$, *then it converges to an integral function $f(z)$ uniformly in every bounded region of the plane, and $f(z)$ satisfies*

$$f^{(2i)}(a) = a_i, \qquad f^{(2i)}(b) = b_i.$$

It is seen from Theorem II that there can be no phenomenon of overconvergence in the Ostrowski sense associated with Lidstone series.

Ostrowski (**5**) showed that, provided that there are gaps in the power series $f(z) = \sum a_n z^n$ satisfying certain conditions, a sequence of partial sums of the power series converges uniformly in a region which extends outside the circle of convergence of $f(z)$.

As is seen from Theorem II, the Lidstone series possesses no region of convergence other than the whole plane, so it is clear that any 'overconvergence' property must be of a different nature. I show that a sequence of partial sums of the Lidstone series may converge to $f(z)$ in the whole plane or in a part of the plane although the Lidstone series itself does not converge for any value of z, except at points congruent to a (mod$(b-a)$). Results of this nature are given in Theorems III, IV, V, and functions are constructed to illustrate the results. I describe the Lidstone series as exhibiting 'overconvergence' in these circumstances.

Before establishing these theorems, we require formulae obtained by Poritsky [see (6) 282–9] in the course of his proof of Theorem I: namely

$$f(z) = \sum_{i=0}^{n} [f^{(2i)}(a)\alpha_i(z) + f^{(2i)}(b)\beta_i(z)] + T_n(z),$$

where
$$T_n(z) = \int_z^a f^{(2n+2)}(s)A_n(z,s)\,ds + \int_z^b f^{(2n+2)}(s)B_n(z,s)\,ds, \tag{6}$$

the integrations being carried out along any paths on which $f(s)$ is analytic, and where

$$\alpha_n(z) = \left.\frac{\partial A_n(z,s)}{\partial s}\right|_{s=a}, \qquad \beta_n(z) = \left.\frac{\partial B_n(z,s)}{\partial s}\right|_{s=b}, \tag{7}$$

$A_n(z,s)$, $B_n(z,s)$ being defined by their generating functions as follows:

$$A(z,s,\lambda) = \sum_{n=0}^{\infty} \lambda^{2n} A_n(z,s) = \frac{\sinh\lambda(z-b)\sinh\lambda(s-a)}{\lambda\sinh\lambda(a-b)}, \tag{8}$$

$$B(z,s,\lambda) = \sum_{n=0}^{\infty} \lambda^{2n} B_n(z,s) = -A(s,z,\lambda). \tag{9}$$

Moreover, if $a = 0$, $b = \pi$, he proves that

$$A_n(z,s),\ -B_n(z,s) = (-1)^n \frac{2}{\pi} \sin z \sin s + O(\tfrac{1}{2}+\epsilon)^{2n} \quad (\epsilon > 0)$$

uniformly for z and s in any bounded part of their planes; and, using elementary transformations, we find that in the general case this becomes

$$A_n(z,s),\ -B_n(z,s)$$
$$= (-1)^n \frac{2(b-a)^{2n+1}}{\pi^{2n+2}} \sin\frac{(z-a)\pi}{b-a} \sin\frac{(s-a)\pi}{b-a} + O\left[\frac{|b-a|}{2\pi} + \epsilon\right]^{2n} \quad (\epsilon > 0). \tag{10}$$

Since (5) implies
$$|f^{(n)}(z)| < Cn^{\frac{1}{2}}k^n \tag{11}$$

if z is restricted to any bounded region [see (6) 289], Poritsky's proof of Theorem I follows. Indeed, condition (5) may be sharpened to

$$f(z) = o\left\{|z|^{-\frac{1}{2}}\exp\pi\left|\frac{z}{a-b}\right|\right\}, \tag{5a}$$

as has been shown by Boas [see (1) Theorem 2 A].

2. Now it follows from (8) and (9) that

$$A(z,s,\lambda)+B(z,s,\lambda) = \frac{\sinh\lambda(s-z)}{\lambda} = \sum_{n=0}^{\infty}\lambda^{2n}\frac{(s-z)^{2n+1}}{(2n+1)!}$$

and hence that

$$A_n(z,s,\lambda)+B_n(z,s,\lambda) = \frac{(s-z)^{2n+1}}{(2n+1)!}.$$

From (6), provided that the integrations are carried out along paths on which $f(z)$ is analytic,

$$\begin{aligned} T_n(z) &= \int_\alpha^a f^{(2n+2)}(s)A_n(z,s)\,ds + \int_\alpha^b f^{(2n+2)}(s)B_n(z,s)\,ds + \\ &\qquad + \int_z^\alpha f^{(2n+2)}(s)\{A_n(z,s)+B_n(z,s)\}\,ds \\ &= \int_\alpha^a f^{(2n+2)}(s)A_n(z,s)\,ds + \int_\alpha^b f^{(2n+2)}(s)B_n(z,s)\,ds + \\ &\qquad + \int_z^\alpha f^{(2n+2)}(s)(s-z)^{2n+1}/(2n+1)!\,ds \\ &= R_n(z)+S_{2n+2}(z,\alpha), \end{aligned}$$

where α is any point, $S_p(z,z_0)$ is the remainder after p terms in the Taylor series for $f(z)$ with centre z_0, and

$$R_n(z) = \int_\alpha^a f^{(2n+2)}(s)A_n(z,s)\,ds + \int_\alpha^b f^{(2n+2)}(s)B_n(z,s)\,ds.$$

Let the circle of convergence of $f(z)$ with centre α be denoted by C_α, and suppose that there exists a rectifiable curve Γ, joining a to b and passing through α, along which $f(z)$ is analytic and on which

$$\overline{\lim_{n\to\infty}}\ \max_{z\in\Gamma}|f^{(2n)}(z)|\{\pi/|a-b|\}^{-2n} = 0. \tag{12}$$

It follows from (10) and (12) that $R_n(z)$ tends to zero uniformly in any

bounded region as n tends to infinity through a suitable sequence. Also $S_{2n+2}(z, \alpha)$ tends to zero uniformly for z in C_α, and we have

THEOREM III. *If $f(z)$ is analytic and satisfies* (12) *on a rectifiable curve Γ joining a and b, then there exists a sequence of partial sums of the Lidstone series* (1) *which converges uniformly to $f(z)$ in the interior of the circle of convergence for $f(z)$ with centre at any point of Γ.*

We have also established

THEOREM IV. *If there exists a sequence n_r such that the corresponding sequence of partial sums S_{2n_r} of the Taylor series for $f(z)$, with centre any point α, converges uniformly to $f(z)$ in a region D and if $f(z)$ is analytic on a rectifiable curve Γ, joining a, α, and b, on which it satisfies*

$$\overline{\lim_{n_r \to \infty}} \max_{z=\Gamma} |f^{(2n_r)}(z)|\{\pi/|a-b|\}^{-2n_r} = 0,$$

then there is a sub-sequence n_{r_p} such that the corresponding sequence of partial sums of the Lidstone series (1) *converges uniformly to $f(z)$ in D.*

3. We also have a theorem about integral functions of 'lower type' less than $\pi/|a-b|$, i.e. an integral function whose maximum modulus satisfies

$$\overline{\lim_{r \to \infty}} \frac{\log M(r)}{r} = \sigma < \frac{\pi}{|a-b|}.$$

THEOREM V. *If $f(z)$ is an integral function satisfying*

$$\overline{\lim_{r \to \infty}} \frac{\log M(r)}{r} = \sigma < \pi/|a-b|,$$

then a sequence of partial sums of the Lidstone series (1) *converges to $f(z)$ for all z, the convergence being uniform in any bounded region.*

This follows from the argument in § 2 with the aid of

LEMMA I. *If $f(z)$ is an integral function satisfying*

$$\overline{\lim_{r \to \infty}} \frac{\log M(r)}{r} = \sigma, \tag{13}$$

then, if D is a bounded set,

$$\overline{\lim_{n \to \infty}} \max_{z \in D} |f^{(2n)}(z)|^{1/2n} \leqslant \sigma.$$

To prove Lemma 1, we observe that, since (13) is satisfied, then, if $l > \sigma$,

$$M(r_n) < \exp l r_n, \tag{14}$$

where r_n is a sequence which tends to infinity. Now

$$f^{(p)}(z) = \frac{p!}{2\pi i} \int_C \frac{f(\zeta)}{(\zeta - z)^{p+1}} d\zeta,$$

where C is any contour surrounding $\zeta = z$. Let C be the contour $|\zeta| = r_n$ where $|z| < r_n$. Then

$$|f^{(p)}(z)| \leqslant \frac{p!M(r_n)r_n}{\{r_n - |z|\}^{p+1}}. \tag{15}$$

Now, if $p_n = [\frac{1}{2}lr_n]$, then, by (14) and (15), and Stirling's approximation for $n!$,

$$|f^{(2p_n)}(z)| \leqslant \frac{(2p_n)!\, e^{2p_n+2}\, l^{2p_n}}{(2p_n)^{2p_n}\{1-\frac{1}{2}|z|l/p_n\}^{2p_n+1}}$$

$$\leqslant \frac{\sqrt{(4\pi p_n)}e^2 l^{2p_n}\{1+o(1)\}}{\{1-\frac{1}{2}|z|l/p_n\}^{2p_n+1}}$$

as n tends to infinity, and therefore

$$|f^{(2p_n)}(z)|^{1/2p_n} \leqslant l\{1+o(1)\}$$

if z is restricted to any bounded set. Therefore

$$\overline{\lim_{n\to\infty}} \max |f^{(2n)}(z)|^{1/2n} \leqslant l,$$

provided that z belongs to any given bounded set. But l is any number greater than σ. The result follows.

4. Functions exist which are not integral functions of the kind considered in Theorem V, but which satisfy the conditions of Theorem III. For example, taking $a = \alpha = 0$, $b = \frac{1}{4}\pi$, Γ the section of the real axis $0 \leqslant x \leqslant \frac{1}{4}\pi$, let

$$\phi(z) = z^{2\lambda_1 - 1} + z^{2\lambda_2 - 1} + \ldots, \tag{16}$$

where λ_n is an increasing sequence of positive integers. The power series (16) has unit radius of convergence and therefore has a singularity at one point (at least) on the unit circle and cannot therefore represent an integral function. The Lidstone series for $\phi(z)$ cannot therefore converge to $\phi(z)$ in any domain D, as in that case (by Theorem II) $\phi(z)$ would coincide with an integral function in D and therefore in the whole plane. We can, however, show that it is possible, provided that x lies on Γ, to select the sequence λ_n in such a way that

$$\max_{0 \leqslant x \leqslant \frac{1}{4}\pi} |\phi^{(2\lambda_n)}(x)| \leqslant 1 \tag{17}$$

and thus, since $\pi/|a-b| = 4$, (12) is satisfied for the sequence λ_n with $f(z) = \phi(z)$, and this function therefore satisfies the conditions of Theorem III.

For, with any integer λ_n and any x on Γ,

$$\frac{d^{2\lambda_n}}{dx^{2\lambda_n}}\left(\frac{1}{1-x}\right) = \sum_{r=2\lambda_n}^{\infty} r(r-1)\ldots(r-2\lambda_n+1)x^{r-2\lambda_n}, \tag{18}$$

and this is a convergent series; thus, if we write

$$\frac{1}{1-x} = \sum_{r=0}^{N-1} x^r + R_N(x),$$

the remainder of the series (18) after $N-2\lambda_n$ terms, namely

$$d^{2\lambda_n} R_N(x)/dx^{2\lambda_n},$$

can be made as small as we please if N is sufficiently large; in fact

$$\frac{d^{2\lambda_n} R_N(x)}{dx^{2\lambda_n}} \leqslant 1$$

provided that $N \geqslant N_0(\lambda_n)$. But, if we choose λ_{n+1} so that

$$N = 2\lambda_{n+1} - 1 \geqslant N_0,$$

then

$$0 \leqslant \phi^{(2\lambda_n)}(x) \leqslant \frac{d^{2\lambda_n} R_N(x)}{dx^{2\lambda_n}},$$

and thus

$$0 \leqslant \phi^{(2\lambda_n)}(x) \leqslant 1,$$

and (17) is satisfied.

We can construct a function to illustrate Theorem IV by a similar treatment of a (suitably chosen) series exhibiting overconvergence in the Ostrowski sense. For example, consider

$$\psi(z) = \sum_{n=0}^{\infty} \{z(1+z)\}^{\mu_n}/p_n, \tag{16 a}$$

where μ_n is an increasing sequence of positive integers satisfying

$$\mu_{n+1} > 2\mu_n + 2,$$

and p_n is the maximal coefficient in the polynomial $\{z(1+z)\}^{\mu_n}$. In each of the polynomials $\{z(1+z)\}^{\mu_n}/p_n$, the moduli of the coefficients do not exceed unity and one of them is actually equal to unity. On the other hand, the degree of the highest term of this polynomial is $2\mu_n$ and that of the lowest term of the next one is μ_{n+1}. Hence, since

$$\mu_{n+1} > 2\mu_n + 2 > 2\mu_n,$$

the individual powers of z in different terms of (16 a) do not overlap, so that the developments of all terms of (16 a) written down successively form a power series $P(z)$ with the radius of convergence 1 (since the moduli of the coefficients of $P(z)$ do not exceed 1 and an infinite number are actually equal to 1). The series (16 a) is the sequence $S_{2\mu_n}$ of partial sums of $P(z)$ and therefore converges uniformly in the interior of the unit-circle.

Noting that $2^{\mu_n}/(\mu_n+1) < p_n < 2^{\mu_n}$, we see that the region of convergence of (16 a) is $|z(1+z)| < 2$: that is a 'dumb-bell shaped' region D

which encloses the circles $|z| < 1$ and $|z+1| < 1$, and touches them at $= -2$ and $z = 1$. Thus $P(z)$ exhibits overconvergence in D in the sense of Ostrowski.

Now, as in the discussion of the function $\phi(z)$, we may make

$$\frac{d^{2\mu_n+2}R_N(z)}{dx^{2\mu_n+2}} \leqslant 1,$$

provided that $N \geqslant N_1(\mu_n)$. So, if we choose μ_{n+1} so that $N = \mu_{n+1} \geqslant N_1$,

$$0 \leqslant \psi^{(2\mu_n+2)}(x) \leqslant \frac{d^{2\mu_n+2}R_N(x)}{dx^{2\mu_n+2}} \leqslant 1,$$

and thus $\psi(z)$ satisfies

$$\max_{0 \leqslant x \leqslant \frac{1}{4}\pi} |\psi^{(2\mu_n+2)}(x)| \leqslant 1.$$

Therefore, if we take $a = \alpha = 0$, $b = \frac{1}{4}\pi$ as before, Γ the segment $0 \leqslant x \leqslant \frac{1}{4}\pi$ of the real axis, $\psi^{(2n)}(x)\{\pi|a-b|\}^{-2n}$ tends to zero as n tends to infinity through the sequence consisting of the integers μ_n+1, and we have seen that the partial sums $S_{2\mu_n+2}$ $(= S_{2\mu_n})$ also converge uniformly in D. Thus all the conditions of Theorem IV are satisfied (n_{r_p} being the complete sequence n_r, identical with the sequence μ_n+1).

II. Gontcharoff Series

5. Poritsky has extended his methods for the Lidstone series to the generalized case where a set of m points replaces the pair of points a, b. Using his formulae, we can prove generalizations of Theorems III, IV, V. It may, however, be of more interest to prove the corresponding theorems for the m-point Gontcharoff series, which has been discussed by Gontcharoff (**3**) and others.

By repeated integration by parts of the multiple integral

$$\overline{T}_{n,m}(z) = \int_{a_1}^{z} dz' \int_{a_2}^{z'} dz'' \ldots \int_{a_n}^{z^{(n-1)}} f^{(n)}(z^{(n)})\, dz^{(n)},$$

where $a_{m+p} = a_p$, we obtain the series in question, viz.

$$\sum_{j=0}^{\infty} \sum_{i=1}^{m} P_{mj+i-1}(z) f^{(mj+i-1)}(a_i), \tag{19}$$

where

$$P_0(z) = 1, \qquad P_k(z) = \int_{a_1}^{z} dz' \int_{a_2}^{z'} dz'' \ldots \int_{a_k}^{z^{(k-1)}} dz^{(k)} \quad (k \geqslant 1).$$

In order to investigate its overconvergence properties I shall find expressions for the remainder $\overline{T}_{n,m}(z)$ in the form of m single integrals.

To be precise, I shall find sequences of functions $\bar{H}_{i,n}(z, x; a_1, a_2, \ldots, a_m)$ to satisfy the identity

$$\int_{a_1}^{z} dz' \int_{a_2}^{z'} dz'' \ldots \int_{a_n}^{z^{(n-1)}} u(x)\, dx = \sum_{i=1}^{m} \int_{a_i}^{z} \bar{H}_{i,n}(z, x; a_1, a_2, \ldots, a_m) u(x)\, dx,$$

where $u(x)$ is an arbitrary analytic function, and thus to satisfy

$$\int_{a_2}^{z'} dz'' \int_{a_3}^{z''} dz''' \ldots \int_{a_n}^{z^{(n-1)}} u(x)\, dx = \sum_{i=1}^{m} \int_{a_{i+1}}^{z'} \bar{H}_{i,n-1}(z', x; a_2, a_3, \ldots, a_m, a_1) u(x)\, dx, \tag{20}$$

where, in both cases $a_{m+p} = a_p$. Clearly

$$\bar{H}_{1,1}(z, x; a_1, a_2, \ldots, a_m) \equiv 1, \qquad \bar{H}_{i,1}(z, x; a_1, a_2, \ldots, a_m) \equiv 0 \quad (i \geqslant 2). \tag{21}$$

Integrating (20) we find

$$\begin{aligned}
&\int_{a_1}^{z} dz' \int_{a_2}^{z'} dz'' \ldots \int_{a_n}^{z^{(n-1)}} u(x)\, dx \\
&\quad = \sum_{i=1}^{m} \int_{a_1}^{z} dz' \int_{a_{i+1}}^{z'} \bar{H}_{i,n-1}(z', x; a_2, a_3, \ldots, a_m, a_1) u(x)\, dx \\
&\quad = \sum_{i=1}^{m-1} \Bigg\{ \int_{a_1}^{z} u(x)\, dx \int_{x}^{a_1} \bar{H}_{i,n-1}(z', x; a_2, a_3, \ldots, a_m, a_1)\, dz' + \\
&\qquad + \int_{a_{i+1}}^{z} u(x)\, dx \int_{a_1}^{z} \bar{H}_{i,n-1}(z', x; a_2, a_3, \ldots, a_m, a_1)\, dz' \Bigg\} + \\
&\qquad + \int_{a_1}^{z} u(x)\, dx \int_{x}^{z} \bar{H}_{m,n-1}(z', x; a_2, a_3, \ldots, a_m, a_1)\, dz'.
\end{aligned}$$

Hence, for all $n \geqslant 2$,

$$\begin{aligned}
\bar{H}_{1,n}(z, x; a_1, a_2, \ldots, a_m) = &\sum_{i=1}^{m-1} \int_{x}^{a_1} \bar{H}_{i,n-1}(z', x; a_2, a_3, \ldots, a_m, a_1)\, dz' + \\
&+ \int_{x}^{z} \bar{H}_{m,n-1}(z', x; a_2, a_3, \ldots, a_m, a_1)\, dz'
\end{aligned} \tag{22}$$

and

$$\bar{H}_{i,n}(z, x; a_1, a_2, \ldots, a_m) = \int_{a_1}^{z} \bar{H}_{i-1,n-1}(z', x; a_2, a_3, \ldots, a_m, a_1)\, dz' \quad (i \geqslant 2) \tag{23}$$

and it follows that

$$\left.\begin{aligned}\frac{\partial \overline{H}_{1,n}(z,x;a_1,a_2,\dots,a_m)}{\partial z} &= \overline{H}_{m,n-1}(z,x;a_2,a_3,\dots,a_m,a_1)\\ \frac{\partial \overline{H}_{i,n}(z,x;a_1,a_2,\dots,a_m)}{\partial z} &= \overline{H}_{i-1,n-1}(z,x;a_2,a_3,\dots,a_m,a_1) \quad (i \geqslant 2)\end{aligned}\right\} \tag{24}$$

and, repeating the process,

$$\frac{\partial^m \overline{H}_{i,n}}{\partial z^m} = \overline{H}_{i,n-m} \quad (\text{all } i\,;\ n \geqslant m+1).$$

By induction, using (21), (22), (23), we have

LEMMA II. $$\sum_{i=1}^{m} \overline{H}_{i,n}(z,x;a_1,a_2,\dots,a_m) = \frac{(z-x)^{n-1}}{(n-1)!}.$$

Also, by (23),

$$\overline{H}_{i,n}(z,x;a_1,a_2,\dots,a_m)|_{z=a_1} = 0 \quad (i \neq 1) \tag{25}$$

and, by (24) and (25),

$$\left.\frac{\partial \overline{H}_{i,n}(z,x;a_1,a_2,\dots,a_m)}{\partial z}\right|_{z=a_2} = 0 \quad (i \neq 2).$$

Repeating the process, we have

$$\left.\frac{\partial^{k-1} \overline{H}_{i,n}(z,x;a_1,a_2,\dots,a_m)}{\partial z^{k-1}}\right|_{z=a_k} = 0 \quad (k = 1,2,\dots,m;\ i \neq k) \tag{26}$$

and hence, by (26) and Lemma II,

$$\left.\begin{aligned}\left.\frac{\partial^{k-1} \overline{H}_{k,n}(z,x;a_1,a_2,\dots,a_m)}{\partial z^{k-1}}\right|_{z=a_k} &= \frac{(a_k-x)^{n-k}}{(n-k)!}\\ &\qquad (k = 1,2,\dots,m;\ n \geqslant k)\\ \left.\frac{\partial^{k-1} \overline{H}_{k,n}(z,x;a_1,a_2,\dots,a_m)}{\partial z^{k-1}}\right|_{z=a_k} &= 0 \quad (n < k)\end{aligned}\right\}. \tag{27}$$

Now let $$\overline{G}_i(z,x;\lambda) = \sum_{n=1}^{\infty} \lambda^{n-1} \overline{H}_{i,n}(z,x;a_1,a_2,\dots,a_m). \tag{28}$$

It follows from (24), (28) that

$$\frac{\partial^m \overline{G}_i}{\partial z^m} = \lambda^m \overline{G}_i. \tag{29}$$

By (26), $$\left.\frac{\partial^{k-1} \overline{G}_i}{\partial z^{k-1}}\right|_{z=a_k} = 0 \quad (k = 1,2,\dots,m;\ i \neq k) \tag{30}$$

and, by (27),

$$\left.\frac{\partial^{k-1} \overline{G}_k}{\partial z^{k-1}}\right|_{z=a_k} = \lambda^{k-1} e^{\lambda(a_k-x)} \quad (k = 1,2,\dots,m). \tag{31}$$

Hence, solving the partial differential equation (29) with boundary conditions (30) and (31), we have

$$\bar{G}_i = A_i e^{\lambda z} + B_i e^{\lambda w z} + \ldots + K_i e^{\lambda w^{m-1} z} \quad \left(w = \exp\frac{2\pi i}{m}\right)$$

where A_i, B_i,..., K_i are functions of x satisfying

$$0 = A_i e^{\lambda a_k} + B_i w^{k-1} e^{\lambda w a_k} + \ldots + K_i w^{(k-1)(m-1)} e^{\lambda w^{m-1} a_k} \quad (k \neq i),$$

$$e^{\lambda(a_i - x)} = A_i e^{\lambda a_i} + B_i w^{i-1} e^{\lambda w a_i} + \ldots + K_i w^{(i-1)(m-1)} e^{\lambda w^{m-1} a_i}.$$

Hence, for example,

$$A_1 = \frac{1}{\bar{D}(\lambda)} \begin{vmatrix} e^{\lambda(a_1 - x)} & e^{\lambda w a_1} & \ldots & e^{\lambda w^{m-1} a_1} \\ 0 & w e^{\lambda w a_2} & \ldots & w^{m-1} e^{\lambda w^{m-1} a_2} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & w^{m-1} e^{\lambda w a_m} & \ldots & w^{(m-1)^2} e^{\lambda w^{m-1} a_m} \end{vmatrix},$$

where

$$\bar{D}(\lambda) = \begin{vmatrix} e^{\lambda a_1} & e^{\lambda w a_1} & \ldots & e^{\lambda w^{m-1} a_1} \\ e^{\lambda a_2} & w e^{\lambda w a_2} & \ldots & w^{m-1} e^{\lambda w^{m-1} a_2} \\ \ldots & \ldots & \ldots & \ldots \\ e^{\lambda a_m} & w^{m-1} e^{\lambda w a_m} & \ldots & w^{(m-1)^2} e^{\lambda w^{m-1} a_m} \end{vmatrix}.$$

Similar formulae may be written down for the remaining functions A_i, B_i,..., K_i. Now it follows from (28) that

$$\bar{H}_{i,n} = \frac{1}{2\pi i} \int_{C_0} \frac{\bar{G}_i(z, x; \lambda)}{\lambda^n} \, d\lambda,$$

where C_0 is a circle $|\lambda| = r < \bar{\rho}$, and $\bar{\rho}$ is the modulus of the root of $\bar{D}(\lambda) = 0$ which is nearest the origin. Now consider the integral

$$\frac{1}{2\pi i} \int_{C_1} \frac{\bar{G}_i(z, x; \lambda)}{\lambda^n} \, d\lambda,$$

where C_1 is a circle $|\lambda| = R > \bar{\rho}$, including only those roots of $\bar{D}(\lambda) = 0$ whose modulus is equal to $\bar{\rho}$. Then

$$\bar{H}_{i,n} = \frac{\bar{F}_i(z, s)}{\bar{\rho}^n} + O(R^{-n}), \tag{32}$$

where $\bar{F}_i$ is a function which is bounded if z and s are in any bounded region. Thus the expansions (28) are valid for $|\lambda| < \bar{\rho}$.

Theorems analogous to Theorems III, IV, V may now be established. For example, the analogue of Theorem III is

THEOREM VI. *If $f(z)$ is analytic and satisfies*

$$\varlimsup_{n\to\infty} \bar{\rho}^{-n} \max_{z\in\Gamma} |f^{(n)}(z)| = 0 \tag{33}$$

uniformly on a rectifiable curve Γ on which $a_1, a_2, \ldots, a_m$ lie, where $\bar{\rho}$ is the modulus of the root nearest the origin of $\overline{D}(\lambda) = 0$, then a sequence of partial sums of the Gontcharoff m-point series (19) *converges uniformly to $f(z)$, in the interior of the circle of convergence for $f(z)$, with centre any point of Γ.*

For the remainder term

$$\overline{T}_{n,m}(z) = \int_{a_1}^{z} dz' \int_{a_2}^{z'} dz'' \ldots \int_{a_n}^{z^{(n-1)}} f^{(n)}(z^{(n)})\, dz^{(n)}$$

$$= \sum_{i=1}^{m} \int_{a_i}^{z} \overline{H}_{i,n}(z, x; a_1, a_2, \ldots, a_m) f^{(n)}(s)\, ds$$

may be written

$$\overline{T}_{n,m}(z) = \sum_{i=1}^{m} \int_{a_i}^{\alpha} \overline{H}_{i,n}(z, s) f^{(n)}(s)\, ds + \int_{\alpha}^{z} \sum_{i=1}^{m} \overline{H}_{i,n}(z, s) f^{(n)}(s)\, ds, \tag{34}$$

and the argument proceeds as for Theorem III, using (32), (33), (34) and Lemma II. It may be noted also that sharpened forms of the classical theorems on the Gontcharoff m-point expansions [see (3)] are now available by the same process as followed by Boas (1) and Gurin (4) for the Lidstone series and its generalization.

6. The case $m = 2$ has been studied by J. M. Whittaker (8), I. J. Schoenberg (7), and others, and it may be of interest to note the special forms taken by some of the above formulae in this case. With $a_{2n+1} = a$, $a_{2n} = b$ (all n) and writing K, $\overline{A}_n(z, x, a, b)$, $\overline{B}_n(z, x, a, b)$ for

$$\pi/2(a-b), \qquad \overline{H}_{1,n}(z, x, a, b), \qquad \overline{H}_{2,n}(z, x, a, b)$$

the remainder of the series after n terms is

$$\int_{a}^{z} \overline{A}_n(z, x, a, b) f^{(n)}(x)\, dx + \int_{b}^{z} \overline{B}_n(z, x, a, b) f^{(n)}(x)\, dx,$$

where

$$\overline{A}_n(z, x, a, b) = \frac{4(-1)^{\frac{1}{2}n}}{\pi K^{n-1}} \cos K(z-b) \sin K(a-x) + O\left(\frac{1}{|3K|} + \epsilon\right)^n \quad (n \text{ even}),$$

$$\overline{A}_n(z, x, a, b) = \frac{4(-1)^{\frac{1}{2}(n-1)}}{\pi K^{n-1}} \cos K(z-b) \cos K(a-x) + O\left(\frac{1}{|3K|} + \epsilon\right)^n$$

(n odd),

$$\bar{B}_n(z,x,a,b) = \frac{4(-1)^{\frac{1}{2}n+1}}{\pi K^{n-1}} \sin K(z-a)\cos K(b-x) + O\left(\frac{1}{|3K|}+\epsilon\right)^n \quad (n \text{ even}),$$

$$\bar{B}_n(z,x,a,b) = \frac{4(-1)^{\frac{1}{2}(n+1)}}{\pi K^{n-1}} \sin K(z-a)\sin K(b-x) + O\left(\frac{1}{|3K|}+\epsilon\right)^n \quad (n \text{ odd}).$$

Also the generating functions (28) reduce to

$$\bar{G}_1(z,x,\lambda) = \frac{e^{\lambda(a-x)}\cosh\lambda(z-b)}{\cosh\lambda(a-b)}$$

and

$$\bar{G}_2(z,x,\lambda) = \frac{e^{\lambda(b-x)}\sinh\lambda(z-a)}{\cosh\lambda(a-b)}.$$

REFERENCES

1. R. P. Boas, Jr., *Duke Math. J.* 10 (1943), 239–45.
2. W. Gontcharoff, *Annales Sci. de l'École Normale* (3), 47 (1930), 1–78.
3. —— *Comm. Soc. Math. Kharkoff* (4) 5 (1932), 67–85.
4. L. S. Gurin, *Mat. Sbornik N.S.* 22 (64) (1948), 425–38.
5. A. Ostrowski, *Berliner Sitzungsb.* (1921) 557–65.
6. H. Poritsky, *Trans. American Math. Soc.* 34 (1932) 274–331.
7. I. J. Schoenberg, ibid. 40 (1936) 12–23.
8. J. M. Whittaker, *Proc. London Math. Soc.* (2) 36 (1933–4), 451–69.

AN ITERATIVE PROCESS FOR THE NUMERICAL DETERMINATION OF CHARACTERISTIC VALUES OF CERTAIN MATRICES

By S. N. AFRIAT (*Oxford*)

[Received 8 February 1950]

1. Introduction

In general terms, when a diagonal element of a Hermitian matrix is sufficiently well separated from the remaining diagonal elements, by the sums of the moduli of non-diagonal elements in respective rows or columns, there is associated with it a characteristic value of the matrix which may be determined indefinitely closely by the iteration to be described. Thus, in particular, when the non-diagonal elements of the matrix are sufficiently small compared with the intervals between the diagonal elements, then all the characteristic values may be so determined.

2. The Process

Let $\mathbf{a}$ be a Hermitian matrix of order n and let a condition R_ν on its elements be given by the inequalities

$$R_\nu\colon |a_{\nu\nu}-a_{\mu\mu}| \geqslant 2\sum_{\kappa\neq\nu}|a_{\nu\kappa}|+\sum_{\kappa\neq\mu}|a_{\mu\kappa}| \quad (\mu\neq\nu).$$

The n circles

$$C_\nu\colon |z-a_{\nu\nu}| \leqslant \sum_{\kappa\neq\nu}|a_{\nu\kappa}| \quad (\nu=1,\ldots,n),$$

which we may call the *diagonal* circles of $\mathbf{a}$, each contain a characteristic value of $\mathbf{a}$ (1); and, if the condition R_ν holds, then the circle C_ν is disjoint from the rest and contains a single characteristic value, say λ_ν, of $\mathbf{a}$ in its interior.

Let $\mathbf{A}_{\nu\nu}$ be the principal square submatrix of $\mathbf{a}$ obtained by omitting the νth row and column; then its $n-1$ diagonal circles $C_{\nu\mu}$ $(\mu\neq\nu)$, containing the characteristic values of $\mathbf{A}_{\nu\nu}$, are each contained in a diagonal circle of $\mathbf{a}$: thus $C_{\nu\mu}$ is contained in C_μ $(\mu\neq\nu)$.

For any λ in C_ν and a characteristic value λ' of $\mathbf{A}_{\nu\nu}$, as a consequence of R_ν we have

$$|\lambda-\lambda'| \geqslant \sum_{\kappa\neq\nu}|a_{\nu\kappa}|.$$

Or, if ω is a characteristic value of $\mathbf{A}_{\nu\nu}-\lambda\mathbf{I}$ for λ in C_ν, then $\lambda+\omega$ is a characteristic value of $\mathbf{A}_{\nu\nu}$ and $|\omega| \geqslant \sum_{\kappa\neq\nu}|a_{\nu\kappa}|$.

Write $\mathbf{a}(\lambda) = \mathbf{a}-\lambda\mathbf{I}, \qquad \mathbf{A}_{\nu\nu}(\lambda) = \mathbf{A}_{\nu\nu}-\lambda\mathbf{I};$

then taking the Cauchy expansion of the determinant $|\mathbf{a}(\lambda)|$ with respect to the νth row and column,

$$|\mathbf{a}(\lambda)| = |\mathbf{A}_{\nu\nu}(\lambda)|\{a_{\nu\nu}-\lambda-\mathbf{a}_{(\nu}\,\mathbf{A}_{\nu\nu}^{-1}(\lambda)\mathbf{a}_{\nu)}\},$$

where $\mathbf{a}_{(\nu}$, $\mathbf{a}_{\nu)}$ are the νth row and column respectively of $\mathbf{a}$ with their νth elements omitted.

Now $|\mathbf{A}_{\nu\nu}(\lambda)| \neq 0$ for λ in C_ν; and, since $|\mathbf{a}(\lambda_\nu)| = 0$, we have

$$a_{\nu\nu}-\lambda_\nu = \mathbf{a}_{(\nu}\,\mathbf{A}_{\nu\nu}^{-1}(\lambda_\nu)\mathbf{a}_{\nu)},$$

$\mathbf{A}_{\nu\nu}^{-1}(\lambda)$ denoting the reciprocal of $\mathbf{A}_{\nu\nu}(\lambda)$.

Further, we have

$$\begin{aligned}\frac{\partial}{\partial\lambda}\mathbf{a}_{(\nu}\,\mathbf{A}_{\nu\nu}^{-1}(\lambda)\mathbf{a}_{\nu)} &= \mathbf{a}_{(\nu}\frac{\partial}{\partial\lambda}\mathbf{A}_{\nu\nu}^{-1}(\lambda)\mathbf{a}_{\nu)}\\ &= \mathbf{a}_{(\nu}\left\{-\mathbf{A}_{\nu\nu}^{-1}(\lambda)\left(\frac{\partial}{\partial\lambda}A_{\nu\nu}(\lambda)\right)\mathbf{A}_{\nu\nu}^{-1}(\lambda)\right\}\mathbf{a}_{\nu)}\\ &= \mathbf{a}_{(\nu}\,\mathbf{A}_{\nu\nu}^{-2}(\lambda)\mathbf{a}_{\nu)}.\end{aligned}$$

Also, if ω is the characteristic value of $\mathbf{A}_{\nu\nu}(\lambda)$ of smallest modulus, then ω^{-2} is the largest characteristic value and so is the bound of $\mathbf{A}_{\nu\nu}^{-2}(\lambda)$. Whence, for λ in C_ν,

$$\begin{aligned}0 \leqslant \mathbf{a}_{(\nu}\,\mathbf{A}_{\nu\nu}^{-2}(\lambda)\mathbf{a}_{\nu)} &\leqslant \omega^{-2}\mathbf{a}_{(\nu}\,\mathbf{a}_{\nu)}\\ &\leqslant \Big(\sum_{\kappa\neq\nu}|a_{\nu\kappa}|^2\Big)\Big(\sum_{\kappa\neq\nu}|a_{\nu\kappa}|\Big)^{-2} < 1,\end{aligned}$$

by the Hermitian character of $\mathbf{a}$, and Schwartz's inequality.

It follows, for the sequence $\lambda_\nu^{(n)}$ $(n = 1, 2,\ldots)$ defined by

$$\lambda_\nu^{(1)} = a_{\nu\nu}, \qquad a_{\nu\nu}-\lambda_\nu^{(n+1)} = \mathbf{a}_{(\nu}\,\mathbf{A}_{\nu\nu}^{-1}(\lambda_\nu^{(n)})\mathbf{a}_{\nu)},$$

that $$\lambda_\nu^{(n)} \to \lambda_\nu \quad (n \to \infty).$$

Thus, whenever a condition R_ν, for some ν, holds for $\mathbf{a}$, a characteristic value λ_ν of $\mathbf{a}$ is calculable to any desired accuracy by computing successive terms of the sequence $\lambda_\nu^{(n)}$. The initial term $\lambda_\nu^{(1)}$ may in fact be anywhere in C_ν, and the error in the nth term is at most

$$\Big\{\Big(\sum_{\kappa\neq\nu}|a_{\nu\kappa}|^2\Big)\Big/\Big(\sum_{\kappa\neq\nu}|a_{\nu\kappa}|\Big)^2\Big\}^n\sum_{\kappa\neq\nu}|a_{\nu\kappa}|.$$

REFERENCE

1. A. Brauer, 'Bounds for the characteristic roots of matrices': *Duke Math. J.* 13 (1946), 387–95.

THE WORD PROBLEM IN A GROUP EXTENSION

By W. H. COCKCROFT (*Aberdeen*)

[Received 27 February 1950]

1. Introduction

LET E be a group extension† of a group Q by a group K: that is to say, there is an invariant subgroup K of E, and a homomorphism $\tau: E \to Q$ mapping E onto Q with kernel K. Suppose that Q and K are given as abstract groups in terms of generators and relations. The main purpose of this note is then to give a set of generators and relations for E and also, under certain conditions, to give a finite construction for the solution of the word problem in E.

An application of the result is given. Let X be a topological space with an Abelian fundamental group γ which has a finite number of generators. Let

$$X^* = X + E_1^2 + \ldots + E_n^2,$$

where E_i^2 $(i = 1, 2, \ldots, n)$ are oriented 2-cells bounded by oriented circuits in X and not meeting X or each other in any point internal to them. The second relative homotopy group $\pi_2(X^*, X)$ of this space is effectively calculated‡ by the methods of the note.

2. Outline of results

Part I. I shall take E (and hence K) as an additive group, although it will not necessarily be commutative, and Q as a multiplicative group. Words in the generators of an additive group will be written as sums, with the symbol 0 for the empty word. In a multiplicative group a word will have its normal meaning in the context, with 1 representing the empty word. Elements of the abstract groups will be regarded as equivalence classes of words w in the ('symbolic') generators, and will be written $\mathbf{w}$ in Part I. I shall use the normal convention of $w_1 \equiv w_2$ for the equivalence of words and $\mathbf{w}_1 = \mathbf{w}_2$ for the equality of elements.

I recall that, by choosing representative elements $\mathbf{u}(\mathbf{x}) \in E$ for all $\mathbf{x} \in Q$, with $\mathbf{u}(\mathbf{x}) \in \tau^{-1}(\mathbf{x})$ and $\mathbf{u}(\mathbf{1}) = \mathbf{0}$, we may define 'normalized' factor-sets $\mathbf{h}(\mathbf{x}, \mathbf{y})$ of Q in K, and automorphisms α of K, by

$$\mathbf{h}(\mathbf{x}, \mathbf{y}) = \mathbf{u}(\mathbf{x}) + \mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{xy}), \qquad \alpha(\mathbf{x})\mathbf{k} = \mathbf{u}(\mathbf{x}) + \mathbf{k} - \mathbf{u}(\mathbf{x})$$

† See (**1**) § 2 for the general definition of a group extension. Compare also the generators and relations defining a group extension in (**2**) § 1 (Theorem 1 in particular).

‡ For a precise definition of *effective calculation* see (**6**) 1237, Note III_4.

for all $\mathbf{x}, \mathbf{y} \subset Q$ and $\mathbf{k} \in K$. We may also define a function f of the generators x_i of Q, with values in E, by setting $f(x_i^\epsilon) = \mathbf{u}(\mathbf{x}_i^\epsilon)$, for all generators x_i of Q, $\epsilon = \pm 1$. This definition may be extended by linearity to words in these generators and, if $R_\lambda \equiv 1$ is any relation of Q, it is trivial that $f(R_\lambda) \in K$.

We assume that definite words are chosen in the classes $\alpha(\mathbf{x}_i^\epsilon)\mathbf{k}_j$ and $f(R_\lambda)$, for all generators x_i of Q and k_j of K, and all relations $R_\lambda \equiv 1$ of Q. These words, which are given in terms of generators k_j of K, will be denoted by $W(\mathbf{x}_i^\epsilon, \mathbf{k}_j)$ and $W(R_\lambda)$, respectively. By means of them we construct in § 3 a *symbolic extension* E^s of Q, with associated homomorphisms τ_s: $E^s \to Q$ and λ_s: $K \approx \tau_s^{-1}(\mathbf{1})$. This extension E^s, which is defined in terms of generators and relations, is proved in § 4 to be equivalent† to E, in the sense that there is an isomorphism $\bar{\theta}$: $E^s \to E$ (onto E) with $\tau\bar{\theta} = \tau_s$, $\bar{\theta}\lambda_s = \lambda$ (λ being the identity isomorphism of K into E).

When K is Abelian, E may be defined‡ in terms of an operator homomorphism, into K, of a free group F, which contains a normal subgroup R such that $F/R \approx Q$. In § 5 I show that in this case the function f can be determined by this operator homomorphism.

In § 6 a solution of the word problem in E^s is given if it is solved in K and (constructively) in Q. By a *constructive* solution of the word problem in a (multiplicative) group is meant a rule for expressing any word which is equivalent to the empty word as a word of the form

$$\prod_{W,R} W R^\epsilon W^{-1},$$

where W is any word in the group, $R \equiv 1$ is any relation, and $\epsilon = \pm 1$.

Part II. The groups Q and K may arise as abstract groups, as above, when E itself is initially given in terms of generators and relations. Thus a representation of Q may then be obtained by adding relations to those of E, and generators and relations of the kernel K of the homomorphism τ: $E \to Q$ may be calculated by the methods of Schreier in (3). The symbolic extension E^s then gives a new representation for E, and, if the word problems can be solved as required in Q and K, we can effectively calculate E.

In Part II below we consider from this point of view the free crossed (γ, d)-module ρ which is an abstract group isomorphic to the group $\pi_2(X^*, X)$ defined in § 1. We shall use the definition of a free crossed

† Compare the definition of (Q, K) equivalence in (1) § 11.
‡ See (5), § 9, Theorem A′.

module given in (4) [§ 2, 455]. The notation used will be that of (4) and the convention of using symbols **w** to denote elements of an abstract group, in order to distinguish them from words w in the generators, will be dropped.

The group ρ is an extension of ρ made Abelian by the commutator subgroup of ρ. Thus effectively to calculate $\pi_2(X^*, X)$

(*a*) we give generators and relations for ρ made Abelian and effectively solve the word problem in it;

(*b*) we give generators and relations for the commutator subgroup of ρ, and solve the word problem in it;

(*c*) we construct the symbolic extension ρ^s which is equivalent to ρ and solve the word problem in it by the methods of § 6.

Finally, as ρ is a group with operators in γ, we calculate the operation of γ on ρ^s.

Part I. The Symbolic Extension

3. Definition of E^s

Let the generators of the group Q be the symbols $(+, x_i)$ with their 'inverses' $(-, x_i)$, written x_i, x_i^{-1} $(i = 1, 2,...)$ and let its relations be $R_\lambda(x_i) \equiv 1$ $(\lambda = 1, 2,...)$. Let the generators of K be the symbols $(+, k_j)$ with their 'negatives' $(-, k_j)$, written k_j, $-k_j$ $(j = 1, 2,...)$, or more briefly ϵk_j $(\epsilon = \pm 1)$. Let its relations be $S_\mu(k_j) \equiv 0$ $(\mu = 1, 2,...)$.

The symbolic extension E^s is then an additive group generated by symbols $(+, X_i)$, $(+, K_j)$, $(-, X_i)$, $(-, K_j)$, written X_i, K_j, $-X_i$, $-K_j$, where the aggregates $\{X_i\}$ and $\{K_j\}$ are in one-to-one correspondence with the aggregates $\{x_i\}$ and $\{k_j\}$ respectively. To obtain the relations of E^s we define a map σ of the generators of K into those of E^s,

$$\sigma(\epsilon k_j) = \epsilon K_j \quad (\epsilon = \pm 1),$$

which is extended by linearity to words in the generators. Let the images, under the map σ, of the words $W(R_\lambda)$ and $W(\mathbf{x}_i^\epsilon, \mathbf{k}_j)$ be f_λ and α_{ij}^ϵ respectively. The relations of E^s are then

$$R_\lambda(X_i) \equiv f_\lambda \qquad \text{(for all } \lambda\text{)}, \tag{1}$$

$$S_\mu(K_j) \equiv 0 \qquad \text{(for all } \mu\text{)}, \tag{2}$$

$$\epsilon X_i + K_j - \epsilon X_i \equiv \alpha_{ij}^\epsilon \qquad \text{(for all } i, j; \epsilon = \pm 1\text{)}, \tag{3}$$

where $R_\lambda(X_i)$ is written additively.

Since $S_\mu(K_j)$, for all μ, are among the relations of E^s, the map σ determines a homomorphism λ_s: $K \to E^s$ (into E^s). Similarly a homomorphism τ_s: $E^s \to Q$ (onto Q) such that $\tau_s^{-1}(1) = \lambda_s K$, is defined by

$$\tau_s(\epsilon \mathbf{X}_i) = \mathbf{x}_i^\epsilon, \qquad \tau_s(\epsilon \mathbf{K}_j) = 1 \quad (\epsilon = \pm 1).$$

4. The equivalence of E^s and E

The extension E may be constructed† as a group of pairs $(\mathbf{k}, \mathbf{x})$ for all $\mathbf{k} \in K$ and $\mathbf{x} \in Q$, with addition defined by

$$(\mathbf{k}, \mathbf{x}) + (\mathbf{m}, \mathbf{y}) = (\mathbf{k} + \alpha(\mathbf{x})\mathbf{m} + \mathbf{h}(\mathbf{x}, \mathbf{y}), \mathbf{xy}),$$

and zero element $(\mathbf{0}, \mathbf{1})$. The correspondence $\mathbf{k} \leftrightarrow (\mathbf{k}, \mathbf{1})$, for all $\mathbf{k} \in K$, defines an isomorphism $\lambda: K \to E$ (into E). Similarly the homomorphism $\tau: E \to Q$ (onto Q) with kernel λK is defined by $\tau(\mathbf{k}, \mathbf{x}) = \mathbf{x}$. The automorphisms of λK are defined by taking $\mathbf{u}(\mathbf{x}) = (\mathbf{0}, \mathbf{x})$, for all $\mathbf{x} \in Q$; thus

$$\alpha(\mathbf{x})(\mathbf{k}, \mathbf{1}) = (\mathbf{0}, \mathbf{x}) + (\mathbf{k}, \mathbf{1}) - (\mathbf{0}, \mathbf{x}).$$

Define a map θ of the generators of E^s to elements of E:

$$\theta(\epsilon X_i) = (\mathbf{0}, \mathbf{x}_i^{\epsilon}); \qquad \theta(\epsilon K_j) = \epsilon(\mathbf{k}_j, \mathbf{1}) \quad \text{(for all } i, j;\ \epsilon = \pm 1).$$

Let θ be extended by linearity to words in the generators of E^s. It may be verified that θ is consistent with the relations of E^s. For instance, suppose that $R_\lambda(X_i) \equiv 0$ is given by

$$\epsilon_1 X_{i_1} + \epsilon_2 X_{i_2} + \ldots + \epsilon_n X_{i_n} \equiv 0 \quad (\epsilon_i = \pm 1).$$

Then we have

$$\begin{aligned} \theta(R_\lambda(X_i)) &= (\mathbf{0}, \mathbf{x}_{i_1}^{\epsilon_1}) + (\mathbf{0}, \mathbf{x}_{i_2}^{\epsilon_2}) + \ldots + (\mathbf{0}, \mathbf{x}_{i_n}^{\epsilon_n}) \\ &= (\mathbf{u}(\mathbf{x}_{i_1}^{\epsilon_1}) + \mathbf{u}(\mathbf{x}_{i_2}^{\epsilon_2}) + \ldots + \mathbf{u}(\mathbf{x}_{i_n}^{\epsilon_n}), \mathbf{x}_{i_1}^{\epsilon_1} \mathbf{x}_{i_2}^{\epsilon_2} \ldots \mathbf{x}_{i_n}^{\epsilon_n}) \\ \text{(for } \mathbf{h}(\mathbf{x}, \mathbf{y}) &= \mathbf{u}(\mathbf{x}) + \mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{xy})) \\ &= (f(R_\lambda), R_\lambda(\mathbf{x}_i)) \\ &= (f(R_\lambda), \mathbf{1}) \\ &= \theta f_\lambda. \end{aligned}$$

Thus θ determines a homomorphism $\bar{\theta}: E^s \to E$ (into E). By definition $\tau\theta = \tau_s$, $\bar{\theta}\lambda_s = \lambda$ and hence $\bar{\theta}: E^s \approx E$, and $\lambda_s^{-1}(0) = 0$ [cf. (7) 435, Lemma 3]. Thus E^s is an extension of Q by K and $\bar{\theta}$ is an equivalence.

5. The symbolic extension when K is Abelian

Let K be Abelian and let Q be regarded‡ as a factor group F/R of a free group F, which is freely generated by the symbols $\{x_i\}$. By Theorem A′ of (5) the extension E may be defined by means of an operator homomorphism of R into K. Indeed, given the factor sets $\mathbf{h}(\mathbf{x}, \mathbf{y})$, this homomorphism is defined as $g|R$ where g is determined, inductively, by the equations

$$\begin{aligned} g(x\,x_i) &= g(x) + \mathbf{h}(\mathbf{x}, \mathbf{x}_i), \\ g(x\,x_i^{-1}) &= g(x) - \mathbf{h}(\mathbf{x}\,\mathbf{x}_i^{-1}, \mathbf{x}_i), \\ g(1) &= \mathbf{0}, \end{aligned}$$

† See (1) § 8.

‡ Compare Kurt Reidemeister, *Einführung in die Kombinatorische Topologie*, Ch. 2, § 5.

where x is any reduced element† of F. Using the definition

$$\mathbf{h}(\mathbf{x}, \mathbf{y}) = \mathbf{u}(\mathbf{x}) + \mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{xy}),$$

it is easily verified that $g|R = f|R$, where f is the homomorphism of F into E defined by $f(x_i) = \epsilon \mathbf{u}(\mathbf{x}_i)$. But, as is readily verified, the function f of § 2 may be defined in this way when K is Abelian, in which case $\theta: E^s \to E$ is altered so that $\theta(\epsilon X_i) = \epsilon(0, x_i)$ and the proof of § 4 remains valid.

Thus, when K is Abelian, the words $W(R_\lambda)$ of § 2 may be defined as representative words of the classes $g(R_\lambda)$, where g: $R \to K$ is the appropriate operator homomorphism associated with E.

6. The word problem in E^s

I define two operations on words in E^s.

Operation (i) Given a word W in E of the form

$$U + \epsilon X_i \pm K_j + V \quad (\epsilon = \pm 1),$$

take the generator ϵX_i to the right of the generator K_j, by using relations (3) of E: thus

$$W \equiv U \pm \alpha_{ij}^{\epsilon} + \epsilon X_i + V;$$

Operation (ii) Suppose that by the reduction processes of Q a word in the generators x_i of Q may be reduced to the empty word and that part of the reduction process is

$$W_1 R_\lambda^{\epsilon} W_2 \to W_1 W_2 \quad (\epsilon = \pm 1).$$

Then given the same word in the generators X_i of E^s we have, by relations (1) and (3) of E^s,

$$W_1 + \epsilon R_\lambda + W_2 \equiv W_1 + \epsilon f_\lambda + W_2 \equiv U + W_1 + W_2,$$

where U is a word in the generators K_j obtained by taking the generators in W_1 to the right of those in f_λ by operations of the type (i). Thus a reduction in Q may be 'copied' in E^s.

Given any word in the generators of E^s, say W, we transform it by operations of the type (i), in a finite number of steps, to the form $W_1 + W_2$, where W_1 is a word in the generators K_j, and W_2 is a word in the generators X_i. If the word W_2 in the generators x_i of Q is not reducible to the empty word, then neither is W_2 in the generators X_i of E^s, for τ_s is a homomorphism. Otherwise, we reduce W_2 in Q and 'copy' the reduction in E^s by operations of the type (ii). W is thus reduced to a word $W_1 + U$ in the generators K_j. By the solution of the word problem in K we know whether such a word is reducible to the empty word by use of relations (2) of E^s. Thus we have a solution of the word problem in E^s.

† i.e. an element in which no formal cancellation of generators and their inverses is possible.

Part II. The Effective Calculation of $\pi_2(X^*, X)$

7. Definition of the free crossed (γ, d)-module ρ

Let γ be an Abelian group with a finite number of generators. Let $\{\alpha_i\}$ $(i = 1, 2, \ldots, n)$ be a finite aggregate of symbols with each of which is associated an element x_i of γ. The free crossed (γ, d)-module ρ is then an additive group whose generators are written (x, α_i), $-(x, \alpha_i)$, for all elements $x \in \gamma$ and all i. Its relations are

$$(x, \alpha_i)+(y, \alpha_j)-(x, \alpha_i) \equiv (x x_i x^{-1} y, \alpha_j) = (x_i y, \alpha_j)$$

for all pairs of elements $x, y \subset \gamma$ and all values of i, j, together with the 'trivial' relations.

It is proved in § 2 of (4) that such a group satisfies the properties of a crossed (γ, d)-module, i.e.

(i) ρ admits γ as a group of operators;

(ii) there is a homomorphism $d\colon \rho \to \gamma$ such that

$$d(xa) = x(da)x^{-1} \quad (a \in \rho,\ x \in \gamma);$$

(iii) $a+b-a = (da)b$, for every pair $a, b \subset \rho$.

I say that, in consequence of the fact that γ is an Abelian group, ρ has the following further properties:

(iv) *the commutator subgroup* $[\rho, \rho]$ *of* ρ *is in the centre of* ρ;

(v) *the commutator* $[a, b]$ *is bilinear in* a *and* b, *and*

$$[-a, b] = [a, -b] = -[a, b]$$

for all pairs $a, b \subset \rho$;

(vi) $[xa, yb] = [a, b]$ *for all* $a, b \subset \rho$, *and all* $x, y \subset \gamma$.

It follows from (iii) that the kernel $d^{-1}(1)$ of the homomorphism $d\colon \rho \to \gamma$ is in the centre of ρ. But

$$d[a, b] = (da)(db)(da)^{-1}(db)^{-1} = 1.$$

Hence $[\rho, \rho] \subset d^{-1}(1)$ and (iv) is proved. We also have

$$\begin{aligned} [a, b+c] &= a+b+c-a-c-b \\ &= (da)b+[a, c]-b, && \text{by (iii)}, \\ &= (da)b-b+[a, c], && \text{by (iv)}, \\ &= [a, b]+[a, c], && \text{by (iii)}, \end{aligned}$$

and

$$\begin{aligned} [a+b, c] &= a+b+c-b-a-c \\ &= a+[b, c]-(dc)a, && \text{by (iii)}, \\ &= a-(dc)a+[b, c], && \text{by (iv)}, \\ &= [a, c]+[b, c], && \text{by (iii)}. \end{aligned}$$

Thus the commutators are bilinear. Also

$$\begin{aligned} [-a,b]+[a,b] &= [0,b], && \text{by (v)},\\ &= 0, \end{aligned}$$

and hence $[-a,b] = -[a,b]$.

Similarly $[a,-b] = -[a,b]$.

Finally we have

$$\begin{aligned} [xa,yb] &= xa+yb-xa-yb\\ &= \{d(xa)\}yb-yb, && \text{by (iii)},\\ &= (da)yb-yb, && \text{by (ii), because } \gamma \text{ is Abelian},\\ &= a+(yb)-a-yb, && \text{by (iii)},\\ &= a-\{d(yb)\}a, && \text{by (iii)},\\ &= a-(db)a, && \text{by (ii)},\\ &= [a,b], && \text{by (iii)}. \end{aligned}$$

Thus (vi) is proved.

It is obvious that ρ is an extension of ρ made Abelian by the commutator subgroup $[\rho,\rho]$. By (iv), ρ is a central extension.

The element of ρ corresponding to the generator $(1,\alpha_i)$ is denoted by a_i, and $da_i = x_i$. Thus any element of ρ is of the form

$$\epsilon_1 y_1 a_{i_1}+\epsilon_2 y_2 a_{i_2}+\ldots+\epsilon_n y_n a_{i_n},$$

where $\epsilon_i = \pm 1$, $y_i \in \gamma$ and repetitions are allowed.

8. The group ρ made Abelian

Let $h\colon \rho \to \rho/[\rho,\rho] = C$ be the homomorphism which makes ρ Abelian. Let $\bar{\gamma}$ denote the factor group $\gamma/d\rho$. The group C is a free $\bar{\gamma}$-module with basis $\{ha_i\}$ [cf. (4) § 2, 457], i.e. if $ha_i = \bar{a}_i$, C is a free Abelian group with generators $\{\bar{x}\bar{a}_i\}$ for all co-sets $\bar{x} \in \bar{\gamma}$ and $i = 1, 2,\ldots, n$.

The group $\bar{\gamma}$, which is obtained by adding relations to those of ρ, is, like γ, an Abelian group with a finite number of generators. Hence the word problem may be solved in $\bar{\gamma}$ and we may calculate the elements $\bar{x} \in \bar{\gamma}$. Thus the generators $\{\bar{x}\bar{a}_i\}$ of C are calculated.

The word problem may obviously be solved in C. Furthermore, it may be solved effectively, for any word equivalent to the empty word is easily expressible as a sum of commutators of pairs of the generators, together with the transforms of such commutators.

9. The commutator subgroup of ρ

It follows from the properties (iv), (v), (vi) of ρ, that the commutator subgroup $[\rho, \rho]$ is the Abelian subgroup of ρ which is generated by the commutators $[a_i, a_j]$. I shall thus denote the generators of $[\rho, \rho]$ by $\pm c_{ij}$ $(i, j = 1, 2,..., n;\ i \neq j)$.

It is obvious that the subgroup $d\rho$ of γ has generating elements $\{x_i\}$ $(i = 1, 2,..., n)$ because γ is Abelian. Let

$$S = x_1^{r_1} x_2^{r_2} ... x_n^{r_n} = 1$$

be any relation among these generating elements. *The group* $[\rho, \rho]$ *is an Abelian group with generators* $\{c_{ij}\}$ *as above, and relations*

$$c_{ij} + c_{ji} \equiv 0 \quad (\textit{for all } c_{ij}),$$

$$\sum_{i=1}^{n} r_i c_{ij} \equiv 0 \quad (\textit{for all relations } S = 1 \textit{ of } d\rho;\ j = 1,..., n;\ i \neq j),$$

together with the trivial relations.

The first relations obviously hold in $[\rho, \rho]$. The second relations also hold, for

$$\sum_{i=1}^{n} r_i [a_i, a_j] = \Big[\sum_{i=1}^{n} r_i a_i, a_j\Big], \quad \text{by (v)},$$

$$= 0$$

because

$$\sum_{i=1}^{n} r_i a_i \in d^{-1}(1),$$

which is in the centre of ρ. We require to prove that these relations form a complete set of relations for $[\rho, \rho]$.

For convenience I shall no longer distinguish between generating elements and symbolic generators of a group. Thus ρ has generators $\{xa_i\}$ for all $x \in \gamma$, $i = 1, 2,..., n$. Add to ρ the generators $\{c_{ij}\}$ and the relations

$$c_{ij} \equiv a_i + a_j - a_i - a_j \quad (i, j = 1, 2,..., n;\ i \neq j).$$

It follows that $[\rho, \rho]$ is now the subgroup of ρ generated by the subset $\{c_{ij}\}$ of the generators of ρ. We thus apply Schreier's theorem [(3) 174, § 3] to prove the completeness of the relations.

First we define a *representative system* $\{u(c)\}$, $u(c) \in h^{-1}(c)$, $u(0) = 0$, for all $c \in C$. By collecting together the generators of C associated with the same basis element and carrying out all possible cancellations, every element c of C may be expressed in a unique form

$$c = \Big(\sum_{j=1}^{p_1} \epsilon_j^{(1)} \bar{y}_j^{(1)} \bar{a}_{i_1}\Big) + ... + \Big(\sum_{j=1}^{p_m} \epsilon_j^{(m)} \bar{y}_j^{(m)} \bar{a}_{i_m}\Big) \quad (\epsilon_j^{(k)} = \pm 1,\ \bar{y}_j^{(k)} \in \bar{\gamma}),$$

where $i_1 < i_2 < ... < i_m$ and the elements $\bar{y}_j^{(k)}$, for each k, are in some

preassigned order given by an ordering of the elements of $\bar{\gamma}$. Thus, taking representatives $v(\bar{y})$ in each co-set $\bar{y}$ of $\bar{\gamma}$, with $v(1) = 1$, we define

$$u(c) = \Big(\sum_{j=1}^{p_1} \epsilon_j^{(1)} v(\bar{y}_j^{(1)}) a_{i_1}\Big) + \ldots + \Big(\sum_{j=1}^{p_m} \epsilon_j^{(m)} v(\bar{y}_j^{(m)}) a_{i_m}\Big).$$

The *Schreier generators* of $[\rho, \rho]$ are then

$$U_{c,xa_i} = u(c) + xa_i - u(c + \bar{x}\bar{a}_i),$$
$$U_{c,c_{ij}} = u(c) + c_{ij} - u(c),$$

for all $c \in C$ and all generators xa_i, c_{ij} of ρ.

Secondly we define a fixed representative word U^*, in the generators c_{ij}, for each Schreier generator U. Associated with the element $c \in C$, defined as above, are the integers

$$l_k = \sum_{j=1}^{p_k} \epsilon_j^{(k)} \quad (k = 1, \ldots, m).$$

It is easily verified that

$$U_{c,xa_i} = xa_i - v(\bar{x})a_i - l_q c_{ii_q} - l_{q+1} c_{ii_{q+1}} - \ldots - l_m c_{ii_m},$$

where $i_{q-1} \leqslant i < i_q$. Define a fixed representative word, in the generators x_i of $d\rho$, for every element $x^{-1}v(\bar{x})$ of $d\rho$. Suppose that $x^{-1}v(\bar{x})$ is represented by

$$x_1^{s_1} x_2^{s_2} \ldots x_n^{s_n} = d(s_1 a_1 + s_2 a_2 + \ldots + s_n a_n).$$

Then

$$U_{c,xa_i} = [s_1 a_1 + \ldots + s_n a_n, xa_i] - \sum_{k=q}^{m} l_k c_{ii_k}$$
$$= \sum_{j=1}^{n} s_j c_{ji} + \sum_{k=q}^{m} l_k c_{i_k i}.$$

Define

$$U^*_{c,c_{ij}} = c_{ij},$$
$$U^*_{c,xa_i} = \sum_{j=1}^{n} s_j c_{ji} + \sum_{k=q}^{m} l_k c_{i_k i}$$

for all generators xa_i, c_{ij} of ρ, and all $c \in C$ defined as above.

Any element $a \in [\rho, \rho]$ can be expressed as a sum of Schreier generators, and on substituting U^* for each generator U, we obtain a word $(a)^*$, say, in the generators c_{ij}. By Schreier's theorem a complete set of relations for $[\rho, \rho]$ is given by

$$(u(c) + R - u(c))^* \equiv 0$$

for all $c \in C$, and all relations $R \equiv 0$ of ρ.

Since $[\rho, \rho]$ is Abelian, these relations will imply

$$c_{ij} + c_{rs} - c_{ij} - c_{rs} \equiv 0,$$

for all c_{ij}, c_{rs}. In what follows I shall assume these commuting relations to be given.

If $R \equiv 0$ is

$$c_{ij}+a_j+a_i-a_j-a_i \equiv 0,$$

then

$$(u(c)+R-u(c))^* = U^*_{c,c_{ij}}+U^*_{c,a_j}+U^*_{c+a_j,a_i}-U^*_{c+a_i,a_j}-U^*_{c,a_i},$$

which is easily verified to be equivalent to

$$c_{ij}+c_{ji} \quad (i>j),$$

or

$$c_{ij}-c_{ij} \quad (i<j),$$

by using the commuting relations of $[\rho, \rho]$.

Again, using the commuting relations of $[\rho, \rho]$, if $R \equiv 0$ is

$$xa_i+ya_j-xa_i-x_iya_j \equiv 0,$$

we have

$$(u(c)+R-u(c))^* = U^*_{c,xa_i}+U^*_{c+\bar{x}\bar{a}_i,ya_j}-U^*_{c+\bar{y}\bar{a}_j,xa_i}-U^*_{c,x_iya_j},$$

which is easily verified to be equivalent to

$$c_{ij}+\sum_{k=1}^{n} p_k c_{kj}-\sum_{k=1}^{n} q_k c_{kj},$$

where $y^{-1}v(\bar{y})$ is represented by $x_1^{p_1}...x_n^{p_n}$ and $x_i^{-1}y^{-1}v(\bar{y})$ is represented by $x_1^{q_1}...x_n^{q_n}$. But these representatives imply that

$$x_1^{p_1-q_1}...x_i^{p_i-q_i+1}...x_n^{p_n-q_n} = 1$$

is a relation of $d\rho$. Moreover,

$$(u(c)+R-u(c))^* \equiv (p_1-q_1)c_{ij}+...+(p_i-q_i+1)c_{ij}+...+(p_k-q_k)c_{kj}.$$

Thus the relations of $[\rho, \rho]$ given above are complete.

The word problem may obviously be solved in $[\rho, \rho]$, for it is an Abelian group with a finite number of generators.

10. The calculation of the equivalent extension

Having calculated C and $[\rho, \rho]$, we can calculate the symbolic extension ρ^s of C by $[\rho, \rho]$, which is equivalent to ρ. Since $[\rho, \rho]$ is Abelian, the extension may be defined in terms of an operator homomorphism. Let F be a free group with generators $\{\bar{x}\bar{a}_i\}$ ($\bar{x} \in \gamma$; $i = 1, 2,..., n$), and let $C \approx F/R$. The operator homomorphism f: $R \to [\rho, \rho]$ which is associated with the extension is defined by $f = g|R$, where g: $F \to \rho$ is given by $g(\bar{x}\bar{a}_i) = \epsilon u(\bar{x}\bar{a}_i)$ ($\epsilon = \pm 1$) [cf. § 5 above].

Since $[\rho, \rho]$ is in the centre of ρ, the inner automorphisms of ρ leave

elements of $[\rho, \rho]$ unaltered. Thus, following the definition of E^s, ρ^s is given by

Generators: C_{ij}, $\bar{X}\bar{A}_i$, for all c_{ij}, $\bar{x}\bar{a}_i$, respectively.

$$\textit{Relations}: \quad \bar{X}\bar{A}_i + \bar{Y}\bar{A}_j - \bar{X}\bar{A}_i - \bar{Y}\bar{A}_j \equiv \left\{ \begin{matrix} C_{ij} & (i \neq j) \\ 0 & (i = j) \end{matrix} \right\}, \tag{1}$$

$$\left.\begin{aligned} &C_{ij} + C_{ji} \equiv 0 \quad \text{(for all } C_{ij}\text{)}, \\ &\sum_{i=1}^{n} r_i C_{ij} \equiv 0 \quad \text{(for all relations } S = 1 \text{ of } d\rho;\ j = 1,\ldots, n;\ i \neq j\text{)}, \\ &C_{ij} + C_{rs} - C_{is} - C_{rj} \equiv 0 \quad \text{(for all } C_{ij}, C_{rs}\text{)}, \end{aligned}\right\} \tag{2}$$

$$\bar{X}\bar{A}_i + C_{ij} - \bar{X}\bar{A}_i \equiv C_{ij} \quad \text{(for all } C_{ij}, \bar{X}\bar{A}_i\text{)}, \tag{3}$$

together with the trivial relations.

The word problem can now be solved in ρ^s by the methods of § 6.

11. The operation of γ on ρ^s

To define the operation of γ on ρ^s, we explicitly define the equivalence ϕ: $\rho \approx \rho^s$, and then define the operation by

$$xa = \phi(x\phi^{-1}(a)) \quad (x \in \gamma,\ a \in \rho^s).$$

The extension ρ may be defined as a group of pairs (k, c), for all $k \in [\rho, \rho]$ and $c \in C$, with the appropriate addition formula, just as E was so defined in § 4 above. In fact ρ may be identified† with this group by an equivalence i defined by

$$i(ku(c)) = (k, c) \quad (k \in [\rho, \rho],\ c \in C),$$

where $u(c) \in h^{-1}(c)$ is a chosen representative of the element $c \in C$. If $\bar{\theta}$ is the equivalence, defined as in § 4 above, between the extension $i\rho$ and the symbolic extension ρ^s, then we define $\phi = \bar{\theta}i$: $\rho \approx \rho^s$.

12. The kernel‡ of d: $\rho \to \gamma$

The kernel of the homomorphism d: $\rho \to \gamma$ can be calculated by Schreier's methods, and ρ can then be constructed as an extension of $d\rho$ by $d^{-1}(1)$. This method is more complicated than the one given above which uses $[\rho, \rho]$ and C. However, we may now calculate $d^{-1}(1)$ directly.

Let the relations among the generating elements $\{x_i\}$ of $d\rho$ be reduced to an independent set

$$x_1^{r_1^{(j)}} x_2^{r_2^{(j)}} \ldots x_n^{r_n^{(j)}} = 1 \quad (j = 1,\ldots, p),$$

the rows of the matrix $[r_i^{(j)}]$ $(i = 1,\ldots, n;\ j = 1,\ldots, p)$ are linearly

† See (**1**) § 8, 334.

‡ $d^{-1}(1) \approx \pi_2(X^*)$ if $\pi_2(X) = 0$.

independent over the ring of integers. Associate with each such relation a symbol δ_j.

Using the chosen representatives $v(\bar{x})$ of the co-sets $\bar{x}$ of $\bar{\gamma}$, associate with each element $v(\bar{x})a_i - a_i$ ($\bar{x} \neq 1$) a symbol $[\bar{x}, a_i]$.

Let A be the free additive Abelian group generated by the symbols δ_j and $[\bar{x}, a_i]$, together with their 'negatives'.

The kernel of d: $\rho \to \gamma$ is isomorphic to the direct sum $A + [\rho, \rho]$.

Define ϕ: $A + [\rho, \rho] \to \rho$ by

$$\phi(\delta_j) = r_1^{(j)}a_1 + \ldots + r_n^{(j)}a_n,$$
$$\phi[\bar{x}, a_i] = v(\bar{x})a_i - a_i,$$
$$\phi(c_{ij}) = [a_i, a_j].$$

It is easily verified that ϕ is a homomorphism into $d^{-1}(1)$. It is an isomorphism, for obviously $\phi \mid [\rho, \rho]$ is an isomorphism, and, if

$$\phi\Big(\sum_i \mu_i \delta_i + \sum_{\bar{x},i} \nu_{\bar{x},i}[\bar{x}, a_i]\Big) = 0,$$

where μ_i, $\nu_{\bar{x},i}$ are integers, then making ρ Abelian, we have

$$\sum_i \sum_{j=1}^{n} \mu_i r_j^{(i)} \bar{a}_j + \sum_{\bar{x},i} \nu_{\bar{x},i} \bar{x}\bar{a}_i - \sum_{\bar{x},i} \nu_{\bar{x},i} \bar{a}_i = 0.$$

Hence $\nu_{\bar{x},i} = 0$ and $\sum_i \mu_i r_j^{(i)} = 0$ ($j = 1, 2, \ldots, n$). But the rows of the matrix $[r_j^{(i)}]$ are independent. Therefore $\mu_i = 0$, for all i.

The isomorphism ϕ is onto $d^{-1}(1)$, for any element of $d^{-1}(1)$, by a simple rearrangement of terms (using the fact that $d^{-1}(1)$ is in the centre of ρ), can be expressed as a sum of the images under ϕ of the generators of A and $[\rho, \rho]$.

I should like to thank Professor J. H. C. Whitehead for his helpful criticism of both exposition and content of this note.

REFERENCES

1. S. Eilenberg and Saunders MacLane, 'Cohomology theory in abstract groups (II)', *Ann. of Math.* 48 (1947), 326–41.
2. O. Schreier, 'Über die Erweiterung von Gruppen' (I), *Monatshefte für Math. und Phys.* (Vienna), 34 (1926), 165–80.
3. ——, 'Die Untergruppen der freien Gruppen', *Abh. Math. Sem. Hansischen Univ.* 5 (1927), 161–83.
4. J. H. C. Whitehead, 'Combinatorial homotopy (II)', *Bull. American Math. Soc.* 55 (1949), 453–96.
5. S. Eilenberg and Saunders MacLane, 'Cohomology theory in abstract groups (III)', *Ann. of Math.* 50 (1949), 736–61.
6. J. H. C. Whitehead, 'On incidence matrices, nuclei, and homotopy types', *Ann. of Math.* 42 (1941), 1197–239.
7. A. L. Blakers, 'Some relations between homology and homotopy groups', *Ann. of Math.* 49 (1948), 428–61.

INEQUALITIES WITH NON-CONJUGATE PARAMETERS

By F. F. BONSALL (*Newcastle-upon-Tyne*)

[Received 30 March 1950]

1. LET

$$p > 1, \quad q > 1, \quad 1/p+1/p' = 1, \quad 1/q+1/q' = 1, \quad 1/p+1/q \geqslant 1,$$

so that $$0 < \lambda = 2-1/p-1/q = 1/p'+1/q' \leqslant 1.$$

Let functions f, g, etc. be non-negative. The parameters p, q are *conjugate* if $1/p+1/q = 1$ (and consequently $\lambda = 1$) and otherwise are *non-conjugate*.

The essence of the present paper is the apparently trivial observation that, since $1/q'+1-\lambda = 1/p$, we have

$$\int kfg\,dx = \int (uf^p)^{1/q'}(vg^q)^{1/p'}(wf^pg^q)^{1-\lambda}\,dx, \tag{1.1}$$

where u, v, w are at our disposal subject only to the condition

$$u^{1/q'}v^{1/p'}w^{1-\lambda} = k.$$

Now, since $1/q'+1/p'+1-\lambda = 1$, Hölder's inequality may be applied to (1.1) and yields

$$\int kfg\,dx \leqslant \Big(\int uf^p\,dx\Big)^{1/q'}\Big(\int vg^q\,dx\Big)^{1/p'}\Big(\int wf^pg^q\,dx\Big)^{1-\lambda}. \tag{1.2}$$

In spite of its trivial appearance, (1.2) is useful for investigating inequalities involving non-conjugate parameters. Results corresponding to (1.2) clearly hold also for multiple integrals and for series.

The technique outlined above is applied in the first place to the extensions and analogues of Hilbert's double series theorem due to Hardy, Littlewood, and Pólya (**1**), Levin (**2**), (**3**) and I. Schur (**4**). In all cases it yields more elegant proofs and in some cases new or sharper results; § 6 contains some remarks on the definition of equi-measurable rearrangements of functions.

Another typical application is a proof of Jensen's extension to non-conjugate parameters of Hölder's inequality for series

$$\sum a_n b_n \leqslant (\sum a_n^p)^{1/p}(\sum b_n^q)^{1/q}. \tag{1.3}$$

In this case a slightly stronger inequality than (1.3) is obtained, and some light is thrown on the domain of validity of the corresponding inequality for integrals.

Quart. J. Math. Oxford (2), 2 (1951), 135–50

In § 11 the technique is generalized to inequalities with n parameters and as an example an extension of Hilbert's theorem with three non-conjugate parameters is obtained.

The remaining sections, § 12 and § 13, are rather loosely connected with the rest. § 12 is concerned with the 'sharpest' form of Hilbert's double series theorem. It is shown that, if the theorem for series is deduced in a perfectly straightforward way from the corresponding theorem for integrals, we obtain

$$\sum_{m,n=1}^{\infty} a_m b_n \left\{ \frac{1}{m+n-1} + \frac{1}{2.3(m+n-1)^3} + \frac{1}{3.5(m+n-1)^5} + \dots \right\} \leqslant \pi \operatorname{cosec} \frac{\pi}{p} \left(\sum_1^{\infty} a_m^p \right)^{1/p} \left(\sum_1^{\infty} b_n^{p'} \right)^{1/p'}.$$

A part of this result is due to Levin (**3**), who, however, obtained it in a much less direct way. A similarly sharpened form of the theorem with non-conjugate parameters is also given. § 13 contains a proof of the known result that $\pi \operatorname{cosec}(\pi/p)$ is the best possible constant in Hilbert's theorem when the parameters are conjugate.

2. Let $f(x) \in L^p(0,\infty)$ and $g(y) \in L^q(0,\infty)$ and write

$$\int_0^{\infty} f^p(x)\,dx = F, \qquad \int_0^{\infty} g^q(y)\,dy = G. \tag{2.1}$$

It was proved by Hardy, Littlewood, and Pólya (**1**) that

$$\int_0^{\infty}\int_0^{\infty} \frac{f(x)g(y)}{(x+y)^{\lambda}}\,dxdy \leqslant K(p,q)F^{1/p}G^{1/q}. \tag{2.2}$$

Their proof, however, threw little light on the value of $K(p,q)$. This situation was improved by Levin (**2**) who proved (2.2) by a method involving only repeated application of Hölder's inequality and obtained the result

$$K(p,q) \leqslant \left(\pi \operatorname{cosec} \frac{\pi}{\lambda p'} \right)^{\lambda}. \tag{2.3}$$

This result, although probably not best-possible, is interesting since it reduces to the known best-possible value when p, q are conjugate. The technique of the present paper enables us to obtain this result by a single application of Hölder's inequality.

THEOREM 1.

$$\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x+y)^\lambda}\,dxdy \leqslant \left(\pi \operatorname{cosec}\frac{\pi}{\lambda p'}\right)^\lambda F^{1/p}G^{1/q}.\dagger \tag{2.4}$$

Proof.

$$I = \int_0^\infty \int_0^\infty \frac{fg}{(x+y)^\lambda}\,dxdy$$

$$= \int_0^\infty \int_0^\infty \left\{\frac{f^p(y/x)^{-t/p'}}{x+y}\right\}^{1/q'} \left\{\frac{g^q(x/y)^{-t/q'}}{x+y}\right\}^{1/p'} \{f^p g^q\}^{1-\lambda}\,dxdy$$

$$\leqslant I_1^{1/q'} I_2^{1/p'} I_3^{1-\lambda}.$$

Here

$$I_3 = \int_0^\infty \int_0^\infty f^p(x)g^q(y)\,dxdy = FG,$$

$$I_1 = \int_0^\infty \int_0^\infty f^p(x)\left(\frac{y}{x}\right)^{-t/p'} \frac{dxdy}{x+y}$$

$$= \int_0^\infty f^p(x)\,dx \int_0^\infty \frac{u^{-t/p'}\,du}{1+u}, \quad \text{on setting } y = xu,$$

$$= \pi \operatorname{cosec}(\pi t/p')F,$$

$$I_2 = \pi \operatorname{cosec}(\pi t/q')G.$$

Combining these results we get

$$I \leqslant \left(\pi \operatorname{cosec}\frac{\pi t}{p'}\right)^{1/q'} \left(\pi \operatorname{cosec}\frac{\pi t}{q'}\right)^{1/p'} F^{1/p}G^{1/q}. \tag{2.5}$$

Now t is at our disposal, and it is a simple matter to show that the least value of the constant here is obtained by taking $t = 1/\lambda$, when (2.5) becomes (2.4). It is interesting that the smallest constant obtained in this way coincides with that obtained by Levin, but there is still no evidence that it is best-possible.

3. In a recent paper (5) Stečkin showed that the series analogue of Theorem 1 is a consequence of the following general lemma:

† All such theorems contain implicitly the statement: if $f \in L^p$ and $g \in L^q$, then the integral on the left exists and satisfies the inequality.

LEMMA. *Let* $\sum\sum a_{mn}x_m y_n$ *and* $\sum\sum b_{mn}x_m y_n$ *be positive bilinear forms with respective bounds* K_1 *and* K_2 *in* $[1/\alpha_1, 1/\beta_1]$ *and* $[1/\alpha_2, 1/\beta_2]$ $(0 < \alpha_j \leqslant 1;\ 0 < \beta_j \leqslant 1)$. *Let* $\alpha = k_1\alpha_1 + k_2\alpha_2$, $\beta = k_1\beta_1 + k_2\beta_2$ *where* $0 < k_j < 1$, $k_1 + k_2 = 1$. *Then* $\sum\sum a_{mn}^{k_1} b_{mn}^{k_2} x_m y_n$ *is bounded in* $[1/\alpha, 1/\beta]$ *with bound* $K_1^{k_1}K_2^{k_2}$.

Stečkin does not consider whether K_1, K_2 being exact bounds implies that $K_1^{k_1}K_2^{k_2}$ is exact. That this is in general untrue may be seen by considering a simple example.

Let $\quad a_{11} = 1, \qquad a_{mn} = 0$ otherwise;

$\qquad b_{11} = b_{12} = 1, \quad b_{mn} = 0$ otherwise.

Then $K_1 = 1$, $K_2 = 2^{1-\beta_2}$, so that $K_1^{k_1}K_2^{k_2} > K$ where $K = 1$ is the bound of $\sum\sum a_{mn}^{k_1} b_{mn}^{k_2} x_m y_n$.

4. THEOREM 2. *Let* $K(x, y)$ *be homogeneous of degree* -1 *and such that*

$$\int_0^\infty u^{-1/\lambda p'} K(1, u)\, du = k,$$

then

$$\int_0^\infty \int_0^\infty f(x)g(y)K^\lambda(x, y)\, dxdy \leqslant k^\lambda F^{1/p} G^{1/q}. \tag{4.1}$$

This is the generalization to the non-conjugate case of a theorem due to I. Schur (4) in the case $p = q = 2$ and extended to the general conjugate case by Hardy, Littlewood, and Pólya [(1) 229].

Proof. As in the proof of Theorem 1,

$$I = \int_0^\infty \int_0^\infty fgK^\lambda\, dxdy \leqslant I_1^{1/q'} I_2^{1/p'} I_3^{1-\lambda},$$

where

$$I_3 = FG,$$

$$I_1 = \int_0^\infty \int_0^\infty f^p(x) \left(\frac{y}{x}\right)^{-1/p'} K(x, y)\, dxdy$$

$$= F \int_0^\infty u^{-1/p'} K(1, u)\, du,$$

$$I_2 = G \int_0^\infty v^{-1/q'} K(v, 1)\, dv.$$

Thus

$$I \leqslant \left\{\int_0^\infty u^{-1/p'} K(1, u)\, du\right\}^{1/q'} \left\{\int_0^\infty v^{-1/q'} K(v, 1)\, dv\right\}^{1/p'} F^{1/p} G^{1/q}. \tag{4.2}$$

The parameter t is at our disposal; and setting $t = 1/\lambda$ and observing that

$$\int_0^\infty u^{-1/\lambda p'} K(1,u)\,du = \int_0^\infty v^{-1/\lambda q'} K(v,1)\,dv = k,$$

we obtain (4.1) from (4.2). This choice of t is the best possible; for, by Hölder's inequality,

$$\left\{\int_0^\infty u^{-t/p'} K(1,u)\,du\right\}^{1/\lambda q'} \left\{\int_0^\infty v^{-t/q'} K(v,1)\,dv\right\}^{1/\lambda p'}$$

$$= \left\{\int_0^\infty u^{-t/p'} K(1,u)\,du\right\}^{1/\lambda q'} \left\{\int_0^\infty u^{t/q'-1} K(1,u)\,du\right\}^{1/\lambda p'}$$

$$\geqslant \int_0^\infty \{u^{-t/p'} K(1,u)\}^{1/\lambda q'} \{u^{t/q'-1} K(1,u)\}^{1/\lambda p'}\,du = k.$$

5. THEOREM 3.

$$\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x+y)^\lambda}\,dxdy \leqslant \mathrm{B}^\lambda\left(\frac{1}{p'},\frac{1}{q'}\right) F^{1/q'} G^{1/p'} \left\{\int_0^\infty \int_0^\infty f^p(x) g^q(y) \frac{x^{1/p'} y^{1/q'}}{(x+y)^\lambda}\,dxdy\right\}^{1-\lambda}.$$

The constant $\mathrm{B}^\lambda = \left\{\dfrac{\Gamma(1/p')\Gamma(1/q')}{\Gamma(1/p'+1/q')}\right\}^\lambda$ *is best-possible.*

This result also implies (2.2), since

$$\frac{x^{1/p'} y^{1/q'}}{(x+y)^\lambda} \leqslant \frac{(1/p')^{1/p'} (1/q')^{1/q'}}{\lambda^\lambda} < 1,$$

but the constant $\left\{\dfrac{(1/p')^{1/p'}(1/q')^{1/q'}}{\lambda^\lambda}\right\}^{1-\lambda} \mathrm{B}^\lambda$

obtained in this way is not in general smaller than that in Theorem 1.

Proof.

$$I = \int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x+y)^\lambda}\,dxdy$$

$$= \int_0^\infty \int_0^\infty \left\{\frac{f^p}{x^{1-\lambda}}\left(\frac{y}{x}\right)^{-1/q} \frac{1}{(x+y)^\lambda}\right\}^{1/q'} \times$$

$$\times \left\{\frac{g^q}{y^{1-\lambda}}\left(\frac{x}{y}\right)^{-1/p} \frac{1}{(x+y)^\lambda}\right\}^{1/p'} \left\{f^p g^q \frac{x^{1/p'} y^{1/q'}}{(x+y)^\lambda}\right\}^{1-\lambda} dxdy$$

$$\leqslant I_1^{1/q'} I_2^{1/p'} I_3^{1-\lambda}.$$

Here

$$I_3 = \int_0^\infty \int_0^\infty f^p(x) g^q(y) \frac{x^{1/p'} y^{1/q'}}{(x+y)^\lambda}\, dxdy,$$

$$I_1 = \int_0^\infty \int_0^\infty \frac{f^p(x)}{x^{1-\lambda}} \left(\frac{y}{x}\right)^{-1/q} \frac{dxdy}{(x+y)^\lambda}$$

$$= \int_0^\infty f^p(x)\, dx \int_0^\infty \frac{u^{-1/q}\, du}{(1+u)^\lambda} = F\mathrm{B}\left(\frac{1}{p'}, \frac{1}{q'}\right),$$

$$I_2 = G\mathrm{B}\left(\frac{1}{p'}, \frac{1}{q'}\right).$$

It remains to prove that B^λ is the best possible constant in this theorem. For this purpose, with $0 < \epsilon < 1$, set

$$f(x) = x^{-1/p} \text{ in } (\epsilon, 1/\epsilon), \qquad f(x) = 0 \text{ elsewhere},$$

$$g(y) = y^{-1/q} \text{ in } (\epsilon, 1/\epsilon), \qquad g(y) = 0 \text{ elsewhere}.$$

Then

$$I = \int_\epsilon^{\epsilon^{-1}} \int_\epsilon^{\epsilon^{-1}} \frac{x^{-1/p} y^{-1/q}}{(x+y)^\lambda} dxdy$$

$$= \int_\epsilon^{\epsilon^{-1}} \frac{dx}{x} \int_\epsilon^{\epsilon^{-1}} \left(\frac{y}{x}\right)^{-1/q} \frac{1}{(1+y/x)^\lambda} \frac{dy}{x}$$

$$= \int_\epsilon^{\epsilon^{-1}} \frac{dx}{x} \int_{\epsilon/x}^{1/\epsilon x} \frac{u^{1/q'-1}}{(1+u)^\lambda} du$$

$$= 2\log(\epsilon^{-1})\mathrm{B} - R_1 - R_2.$$

$$R_1 = \int_\epsilon^{\epsilon^{-1}} \frac{dx}{x} \int_0^{\epsilon/x} \frac{u^{1/q'-1}}{(1+u)^\lambda} du < \int_\epsilon^{\epsilon^{-1}} \frac{dx}{x} \int_0^{\epsilon/x} u^{1/q'-1}\, du = q'^2(1-\epsilon^{2/q'}) = O(1),$$

$$R_2 = \int_\epsilon^{\epsilon^{-1}} \frac{dx}{x} \int_{1/\epsilon x}^\infty \frac{u^{1/q'-1}}{(1+u)^\lambda} du < \int_\epsilon^{\epsilon^{-1}} \frac{dx}{x} \int_{1/\epsilon x}^\infty u^{-1/p'-1}\, du = p'^2(1-\epsilon^{2/p'}) = O(1).$$

Thus

$$I = 2\log(1/\epsilon)\mathrm{B} + O(1),$$

$$F = \int_0^\infty f^p(x)\, dx = \int_\epsilon^{\epsilon^{-1}} \frac{dx}{x} = 2\log(1/\epsilon), \qquad G = 2\log(1/\epsilon),$$

$$I_3 = \int_\epsilon^{\epsilon^{-1}} \int_\epsilon^{\epsilon^{-1}} \frac{x^{-1/p} y^{-1/q}}{(x+y)^\lambda} dxdy = 2\log(1/\epsilon)\mathrm{B} + O(1).$$

It follows that

$$B^{\lambda}F^{1/q'}G^{1/p'}I_3^{1-\lambda} = 2\log(\epsilon^{-1})B + o\{\log(\epsilon^{-1})\},$$

and the theorem is now completed by letting $\epsilon \to 0$.

6. In the following sections, use is made of the equi-measurable rearrangement of a function in decreasing order. Let $\phi(x)$ be a measurable non-negative function of $L(0,\infty)$; then the measure $M(y)$ of the set in which $\phi(x) \geqslant y$ is finite for $y > 0$ and decreasing. Hardy, Littlewood, and Pólya [(1) 276] define the decreasing rearrangement $\bar{\phi}(x)$ of the function $\phi(x)$ by the equation

$$\bar{\phi}\{M(y)\} = y$$

and remark that this definition is not quite complete. In fact, if $\phi(x)$ is a step function, $\bar{\phi}(x)$ is not defined anywhere by this equation. This difficulty may be avoided by making the definition

$$\bar{\phi}(x) = \inf_{M(y)\leqslant x} y. \tag{6.1}$$

If then $\bar{M}(c)$ denotes the measure of the set in which $\bar{\phi}(x) \geqslant c$, it can be proved by well-known methods that

$$\bar{M}(c) = M(c).$$

Consequently the measure of the set in which $c_1 \leqslant \bar{\phi}(x) < c_2$ is the same as the measure of the set in which $c_1 \leqslant \phi(x) < c_2$.

7. It was proved by Levin (2) that, if $\bar{f}$, $\bar{g}$ are the equi-measurable rearrangements of f, g in decreasing order, then

$$\int_0^\infty\int_0^\infty \frac{f(x)g(y)}{(x+y)^{\lambda}}\,dxdy \leqslant \int_0^\infty\int_0^\infty \frac{\bar{f}(x)\bar{g}(y)}{(x+y)^{\lambda}}\,dxdy. \tag{7.1}$$

Setting $$\phi = \sup_{0<x<\infty} x\bar{f}^{p}(x), \qquad \psi = \sup_{0<y<\infty} y\bar{g}^{q}(y),$$

Levin deduced that

$$\int_0^\infty\int_0^\infty \frac{f(x)g(y)}{(x+y)^{\lambda}}\,dxdy \leqslant k\phi^{(1-\lambda)/\lambda p'}\psi^{(1-\lambda)/\lambda q'}F^{1/\lambda q'}G^{1/\lambda p'} \tag{7.2}$$

and that $B(1/p', 1/q')$ is the best possible value of k.

Writing $$\alpha = \frac{\phi}{F}, \qquad \beta = \frac{\psi}{G}, \tag{7.3}$$

we see that (7.2) is equivalent to

$$\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x+y)^\lambda}\,dxdy \leqslant \mathrm{B}\left(\frac{1}{p'}, \frac{1}{q'}\right) \alpha^{(1-\lambda)/\lambda p'} \beta^{(1-\lambda)/\lambda q'} F^{1/p} G^{1/q}. \tag{7.4}$$

Now although the constant B in (7.4) is exact, it is nevertheless possible to assert a sharper inequality of this type.

THEOREM 4. *Let γ be the smaller of α, β defined by* (7.3). *Then*

$$\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x+y)^\lambda}\,dxdy \leqslant \mathrm{B}\left(\frac{1}{p'}, \frac{1}{q'}\right) \gamma^{1-\lambda} F^{1/p} G^{1/q},$$

and the constant $\mathrm{B}(1/p', 1/q')$ is best-possible.

This inequality is evidently sharper than (7.4) except in the special case when $\alpha = \beta$; the constant B is thus best-possible. It is an interesting feature of this inequality that the term γ depends on only one of the functions f, g.

Proof. Suppose, for example, that $\gamma = \alpha \leqslant \beta$; then

$$\begin{aligned}
&\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x+y)^\lambda}\,dxdy \\
&\qquad \leqslant \int_0^\infty \int_0^\infty \frac{\bar{f}(x)\bar{g}(y)}{(x+y)^\lambda}\,dxdy, \quad \text{by (7.1)}, \\
&\qquad \leqslant \mathrm{B}^\lambda F^{1/q'} G^{1/p'} \left\{ \int_0^\infty \int_0^\infty \frac{\bar{f}^p \bar{g}^q x^{1/p'} y^{1/q'}}{(x+y)^\lambda}\,dxdy \right\}^{1-\lambda}, \quad \text{by Theorem 3}, \\
&\qquad = \mathrm{B}^\lambda F^{1/q'} G^{1/p'} \left\{ \int_0^\infty \bar{g}^q(y)\,dy \int_0^\infty [x\bar{f}^p(x)] \left(\frac{x}{y}\right)^{1/p'-1} \frac{1}{(1+x/y)^\lambda} \frac{dx}{y} \right\}^{1-\lambda} \\
&\qquad \leqslant \mathrm{B}^\lambda \alpha^{1-\lambda} F^{1/p} G^{1/p'} \left\{ \int_0^\infty \bar{g}^q(y)\,dy \int_0^\infty \frac{u^{1/p'-1}}{(1+u)^\lambda}\,du \right\}^{1-\lambda} \\
&\qquad = \mathrm{B}\alpha^{1-\lambda} F^{1/p} G^{1/q}.
\end{aligned}$$

8. The method of Theorems 3 and 4 can be applied to the double integral

$$\int_{-\infty}^\infty \int_{-\infty}^\infty \frac{f(x)g(y)}{|x-y|^\lambda}\,dxdy \quad (0 < \lambda < 1). \tag{8.1}$$

Hardy, Littlewood, and Pólya [(1), Chapter X] prove that the boundedness of this bilinear form is a consequence of the boundedness of

$$\int_0^\infty \int_0^\infty \frac{f^*(x)g^*(y)}{|x-y|^\lambda}\,dxdy, \tag{8.2}$$

where f^*, g^* are restricted to be non-increasing functions. This latter result is then obtained in a somewhat roundabout manner. The following theorem establishes the boundedness of (8.2) for all f^* and all non-increasing g^*. It also gives an estimate for the bound.

THEOREM 5. *If $g(y)$ is a non-increasing function and $0 < \lambda < 1$, then*

$$\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{|x-y|^\lambda}\,dxdy \leqslant K F^{1/p} G^{1/q}, \tag{8.3}$$

where

$$K \leqslant \beta^{1-\lambda}\Gamma(1-\lambda)\left\{\frac{\Gamma(1/p')}{\Gamma(1/q)} + \frac{\Gamma(1/q')}{\Gamma(1/p)}\right\}. \tag{8.4}$$

Since

$$\beta = \frac{\sup yg^q(y)}{G} \leqslant 1,$$

the theorem implies that

$$K \leqslant \Gamma(1-\lambda)\left\{\frac{\Gamma(1/p')}{\Gamma(1/q)} + \frac{\Gamma(1/q')}{\Gamma(1/p)}\right\} = M, \text{ say.} \tag{8.5}$$

Proof.

$$I = \int_0^\infty \int_0^\infty \frac{f(x)g(y)}{|x-y|^\lambda}\,dxdy \leqslant I_1^{1/q'} I_2^{1/p'} I_3^{1-\lambda},$$

where

$$\begin{aligned}
I_1 &= \int_0^\infty \int_0^\infty \frac{f^p(x)}{x^{1-\lambda}} \left(\frac{y}{x}\right)^{-1/q} \frac{dxdy}{|x-y|^\lambda} \\
&= \int_0^\infty f^p(x)\,dx \int_0^\infty \frac{u^{-1/q}}{|1-u|^\lambda}\,du \\
&= F\left[\int_0^1 \frac{u^{-1/q}\,du}{(1-u)^\lambda} + \int_1^\infty \frac{u^{-1/q}}{(u-1)^\lambda}\,du\right] \\
&= F\left[\int_0^1 u^{1/q'-1}(1-u)^{-\lambda}\,du + \int_0^1 v^{1/p'-1}(1-v)^{-\lambda}\,dv\right] \\
&= F\left[\mathrm{B}\left(\frac{1}{q'}, 1-\lambda\right) + \mathrm{B}\left(\frac{1}{p'}, 1-\lambda\right)\right] \\
&= FM.
\end{aligned}$$

Similarly

$$I_2 = GM$$

$$I_3 = \int_0^\infty \int_0^\infty f^p(x) g^q(y) \frac{x^{1/p'} y^{1/q'}}{|x-y|^\lambda}\, dxdy$$

$$\leqslant \beta G \int_0^\infty f^p(x)\, dx \int_0^\infty \frac{x^{1/p'} y^{1/q'-1}}{|x-y|^\lambda}\, dy$$

$$= \beta G \int_0^\infty f^p(x)\, dx \int_0^\infty \frac{u^{1/q'-1}}{|1-u|^\lambda}\, du$$

$$= \beta F G M.$$

Combining these results, we obtain (8.3) and (8.4).

9. Hölder's inequality for series was generalized by Jensen [(1) 29] in the following manner:

If $\alpha, \beta, \ldots, \lambda$ are positive and $\alpha+\beta+\ldots+\lambda \geqslant 1$, then

$$\sum a^\alpha b^\beta \ldots l^\lambda \leqslant (\sum a)^\alpha (\sum b)^\beta \ldots (\sum l)^\lambda. \tag{9.1}$$

This includes as a special case

$$\sum ab \leqslant (\sum a^p)^{1/p} (\sum b^q)^{1/q}, \tag{9.2}$$

where $p > 1$, $q > 1$ and $p^{-1}+q^{-1} \geqslant 1$. The inequality (9.2) and indeed sharper inequalities of the same type can be proved by the methods of the present paper.

THEOREM 6. *Let $p > 1$, $q > 1$, $p^{-1}+q^{-1} \geqslant 1$. Let the sequences $(\bar{a}_n)$, $(\bar{b}_n)$ be the rearrangements of (a_n), (b_n) in decreasing order,*

$$\alpha = \sup \frac{n\bar{a}_n^p}{\sum_1^\infty a_n^p}, \qquad \beta = \sup \frac{n\bar{b}_n^q}{\sum_1^\infty b_n^q}$$

and let γ be the smaller of α, β. Then

$$\text{(i)} \quad \sum_1^\infty a_n b_n \leqslant \alpha^{1-\lambda} \Big(\sum_1^\infty a_n^p\Big)^{1/p} \Big(\sum_1^\infty b_n^q\Big)^{1/p'} \Big(\sum_1^\infty \frac{\bar{b}_n^q}{n}\Big)^{1-\lambda},$$

$$\text{(ii)} \quad \sum_1^\infty a_n b_n \leqslant \gamma^{1-\lambda} \Big(\sum_1^\infty a_n^p\Big)^{1/p} \Big(\sum_1^\infty b_n^q\Big)^{1/q},$$

$$\text{(iii)} \quad \sum_1^\infty a_n b_n \leqslant \{\tfrac{1}{6}\pi^2 \alpha\beta\}^{1-\lambda} \Big(\sum_1^\infty a_n^p\Big)^{1/p} \Big(\sum_1^\infty b_n^q\Big)^{1/q}.$$

Since $\gamma \leqslant 1$, the inequality (ii) is evidently sharper than (9.2).

Proof. Suppose that $\gamma = \alpha \leqslant \beta$; then

$$\sum_1^\infty a_n b_n = \sum_1^\infty (a_n^p)^{1/q'} (b_n^q)^{1/p'} (a_n^p b_n^q)^{1-\lambda}$$

$$\leqslant \Big(\sum_1^\infty a_n^p\Big)^{1/q'} \Big(\sum_1^\infty b_n^q\Big)^{1/p'} S^{1-\lambda},$$

where $$S = \sum_1^\infty a_n^p b_n^q \leqslant \sum_1^\infty \bar{a}_n^p \bar{b}_n^q \leqslant \alpha \sum_1^\infty \bar{a}_n^p \sum_1^\infty \frac{\bar{b}_n^q}{n}.$$

This proves (i) and, since

$$\sum_1^\infty \frac{\bar{b}_n^q}{n} \leqslant \sum_1^\infty b_n^q,$$

(ii) is an obvious corollary. The inequality (iii) follows from (i) since

$$\sum_1^\infty \frac{\bar{b}_n^q}{n} \leqslant \beta \sum_1^\infty b_n^q \sum_1^\infty \frac{1}{n^2} = \tfrac{1}{6}\pi^2 \beta \sum_1^\infty b_n^q.$$

10. Owing to the lack of 'homogeneity in $\sum$' the inequality for integrals corresponding to (9.2) is not valid. The following theorem is, however, in some respects analogous to (9.2).

THEOREM 7. *If*

$$F = \int_a^b f^p(x)\,dx, \qquad G = \int_a^b g^q(x)\,dx, \qquad \int_a^b f^p(x) g^q(x)\,dx \leqslant \kappa F G,$$

then

$$\int_a^b f(x) g(x)\,dx \leqslant \kappa^{1-\lambda} F^{1/p} G^{1/q}.$$

Here p, q, λ satisfy the usual conditions and the proof of the theorem is an obvious application of the methods of this paper. A particular case of this theorem is of some interest. It is known [(**1**) 168] that, if f, g are oppositely ordered functions, i.e. $\{f(x_1)-f(x_2)\}\{g(x_1)-g(x_2)\} \leqslant 0$ for all x_1, x_2 in (a, b), then

$$\int_a^b dx \int_a^b f^p(x) g^q(x)\,dx \leqslant \int_a^b f^p(x)\,dx \int_a^b g^q(x)\,dx. \tag{10.1}$$

This leads at once to the following corollary of Theorem 7.

COROLLARY. *If f, g are oppositely ordered functions in (a, b), then*

$$\int_a^b f(x)g(x)\,dx \leqslant \frac{1}{(b-a)^{1-\lambda}} F^{1/p} G^{1/q}. \tag{10.2}$$

11. The present method of proving inequalities with non-conjugate parameters may be generalized to the case of n such parameters.

Let $i = 1, 2, \ldots, n$ and let p_i, p_i', q_i, λ satisfy

$$\left.\begin{aligned} &(a)\quad p_i > 1, \quad 1/p_i + 1/p_i' = 1, \\ &(b)\quad \sum_i 1/p_i \geqslant 1, \\ &(c)\quad \lambda = \frac{1}{n-1} \sum_i 1/p_i', \\ &(d)\quad 1/q_i = \lambda - 1/p_i' \geqslant 0. \end{aligned}\right\} \tag{11.1}$$

It follows from these conditions that

$$\frac{1}{q_1} + \frac{1}{q_2} + \ldots + \frac{1}{q_n} + 1 - \lambda = 1,$$

$$\frac{1}{q_i} + 1 - \lambda = 1 - \frac{1}{p_i'} = \frac{1}{p_i}, \qquad \lambda \leqslant 1.$$

Hence,

$$\int k f_1 f_2 \ldots f_n\,dx = \int (u_1 f_1^{p_1})^{1/q_1} \ldots (u_n f_n^{p_n})^{1/q_n} (w f_1^{p_1} f_2^{p_2} \ldots f_n^{p_n})^{1-\lambda}\,dx$$

$$\leqslant \left(\int u_1 f_1^{p_1}\,dx\right)^{1/q_1} \ldots \left(\int u_n f_n^{p_n}\,dx\right)^{1/q_n} \left(\int w f_1^{p_1} \ldots f_n^{p_n}\,dx\right)^{1-\lambda}, \tag{11.2}$$

where $u_1, \ldots, u_n$, w are at our disposal subject only to the condition

$$u_1^{1/q_1} u_2^{1/q_2} \ldots u_n^{1/q_n} w^{1-\lambda} = k.$$

To illustrate this extension of the method, I prove Hilbert's inequality with three non-conjugate parameters.

THEOREM 8. *Let p_1, p_2, p_3 satisfy conditions* (11.1) *with $n = 3$; let*

$$F = \int_0^\infty f^{p_1}(x)\,dx, \qquad G = \int_0^\infty g^{p_2}(y)\,dy, \qquad H = \int_0^\infty h^{p_3}(z)\,dz.$$

Then

$$\int_0^\infty \int_0^\infty \int_0^\infty \frac{f(x)g(y)h(z)}{(x+y+z)^{2\lambda}}\,dx\,dy\,dz \leqslant k F^{1/p_1} G^{1/p_2} H^{1/p_3},$$

where

$$k \leqslant \left\{\Gamma\left(\frac{1}{\lambda p_1'}\right) \Gamma\left(\frac{1}{\lambda p_2'}\right) \Gamma\left(\frac{1}{\lambda p_3'}\right)\right\}^\lambda.$$

Proof. With α, β, γ at our disposal, we may write

$$I = \int_0^\infty\int_0^\infty\int_0^\infty \frac{f(x)g(y)h(z)}{(x+y+z)^{2\lambda}}\,dxdydz$$

$$= \int_0^\infty\int_0^\infty\int_0^\infty \left\{\frac{f^{p_1}(x)(y/x)^{-\gamma q_1}(z/x)^{-\beta q_1}}{(x+y+z)^2}\right\}^{1/q_1}\left\{\frac{g^{p_2}(y)(z/y)^{-\alpha q_2}(x/y)^{-\gamma q_2}}{(x+y+z)^2}\right\}^{1/q_2}\times$$

$$\times\left\{\frac{h^{p_3}(z)(x/z)^{-\beta q_3}(y/z)^{-\alpha q_3}}{(x+y+z)^2}\right\}^{1/q_3}\{f^{p_1}(x)g^{p_2}(y)h^{p_3}(z)\}^{1-\lambda}\,dxdydz$$

$$\leqslant I_1^{1/q_1}I_2^{1/q_2}I_3^{1/q_3}I_4^{1-\lambda},$$

$$I_4 = \int_0^\infty\int_0^\infty\int_0^\infty f^{p_1}(x)g^{p_2}(y)h^{p_3}(z)\,dxdydz = FGH,$$

$$I_1 = \int_0^\infty\int_0^\infty\int_0^\infty \frac{f^{p_1}(x)(y/x)^{-\gamma q_1}(z/x)^{-\beta q_1}}{(x+y+z)^2}\,dxdydz$$

$$= \int_0^\infty f^{p_1}(x)\,du\int_0^\infty\int_0^\infty \frac{u^{-\gamma q_1}v^{-\beta q_1}}{(1+u+v)^2}\,dudv = FJ_1, \text{ say.}$$

Setting $u+v=\xi$, $v=\xi\eta$, we find

$$J_1 = \int_0^\infty d\xi\int_0^1 \frac{\xi^{-\gamma q_1}(1-\eta)^{-\gamma q_1}\xi^{-\beta q_1}\eta^{-\beta q_1}}{(1+\xi)^2}\,\xi\,d\eta$$

$$= \int_0^\infty \frac{\xi^{1-\beta q_1-\gamma q_1}}{(1+\xi)^2}\,d\xi\int_0^1 \eta^{-\beta q_1}(1-\eta)^{-\gamma q_1}\,d\eta$$

$$= \mathrm{B}(2-\beta q_1-\gamma q_1, \beta q_1+\gamma q_1)\mathrm{B}(1-\beta q_1, 1-\gamma q_1)$$

$$= \Gamma(\beta q_1+\gamma q_1)\Gamma(1-\beta q_1)\Gamma(1-\gamma q_1).$$

Similarly $I_2 = GJ_2$ and $I_3 = HJ_3$, where

$$J_2 = \Gamma(\gamma q_2+\alpha q_2)\Gamma(1-\gamma q_2)\Gamma(1-\alpha q_2),$$

$$J_3 = \Gamma(\alpha q_3+\beta q_3)\Gamma(1-\alpha q_3)\Gamma(1-\beta q_3).$$

Now α, β, γ are at our disposal, and setting

$$\alpha = \frac{1}{\lambda q_2 q_3}, \qquad \beta = \frac{1}{\lambda q_3 q_1}, \qquad \gamma = \frac{1}{\lambda q_1 q_2}$$

and observing that

$$\frac{1}{\lambda}\left(\frac{1}{q_2}+\frac{1}{q_3}\right) = \frac{1}{\lambda p_1'}, \qquad 1-\frac{1}{\lambda q_3} = \frac{1}{\lambda p_3'}, \quad \text{etc.},$$

we obtain

$$J_1 = J_2 = J_3 = \Gamma\left(\frac{1}{\lambda p_1'}\right)\Gamma\left(\frac{1}{\lambda p_2'}\right)\Gamma\left(\frac{1}{\lambda p_3'}\right) = J, \text{ say.}$$

We have now proved that

$$I \leqslant (J_1 F)^{1/q_1}(J_2 G)^{1/q_2}(J_3 H)^{1/q_3}(FGH)^{1-\lambda} = J^{\lambda} F^{1/p_1} G^{1/p_2} H^{1/p_3}.$$

12. It is well known [(**1**) 234] that Hilbert's theorem for series can be stated in the sharpened form

$$\sum_{m,n=1}^{\infty} \frac{a_m b_n}{m+n-1} \leqslant \pi \operatorname{cosec}\frac{\pi}{p}\left(\sum_1^{\infty} a_m^p\right)^{1/p}\left(\sum_1^{\infty} b_n^{p'}\right)^{1/p'}. \tag{12.1}$$

In the special case $p = p' = 2$, it was proved by Levin (**3**) that the kernel $1/(m+n-1)$ on the left-hand side can be replaced by the larger kernel

$$\log\frac{(m+n-2)^{m+n-2}(m+n)^{m+n}}{(m+n-1)^{2(m+n-1)}}, \tag{12.2}$$

without other change of the inequality. Levin's proof, however, is complicated and does not reveal the simple nature of this result: *that it is a direct consequence of the truth of Hilbert's theorem for integrals.*

By deducing Levin's result from the integral theorem we can obtain it with complete generality, i.e. for all p in the conjugate case and a similar sharpening in the non-conjugate case. The result appears in a more interesting form if we expand (12.2) in descending powers of $m+n-1$, when it becomes evident that (12.2) is greater than

$$1/(m+n-1).$$

In fact, writing $m+n-1 = t$, we find

$$\log\frac{(m+n-2)^{m+n-2}(m+n)^{m+n}}{(m+n-1)^{2(m+n-1)}} = (t-1)\log(1-1/t)+(t+1)\log(1+1/t)$$

$$= \frac{1}{t}+\frac{1}{2.3t^3}+\frac{1}{3.5t^5}+\ldots. \tag{12.3}$$

THEOREM 9.

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_m b_n\left\{\frac{1}{m+n-1}+\frac{1}{2.3(m+n-1)^3}+\frac{1}{3.5(m+n-1)^5}+\ldots\right\}$$

$$\leqslant \pi \operatorname{cosec}\frac{\pi}{p}\left(\sum_1^{\infty} a_m^p\right)^{1/p}\left(\sum_1^{\infty} b_n^{p'}\right)^{1/p'}.$$

Proof. It is well-known (case $\lambda = 1$ of Theorem 1) that

$$\int_0^\infty \int_0^\infty \frac{f(x)g(y)}{(x+y)}\,dxdy \leqslant \pi \operatorname{cosec}\frac{\pi}{p}\left\{\int_0^\infty f^p(x)\,dx\right\}^{1/p}\left\{\int_0^\infty g^{p'}(y)\,dy\right\}^{1/p'}. \tag{12.4}$$

In (12.4) let us make the specialization

$$\left.\begin{aligned} f(x) &= a_m \quad (m-1 \leqslant x < m) \\ g(y) &= b_n \quad (n-1 \leqslant y < n) \end{aligned}\right\}. \tag{12.5}$$

Then

$$\sum_{m=1}^\infty \sum_{n=1}^\infty K_{m,n}\,a_m b_n \leqslant \pi \operatorname{cosec}\frac{\pi}{p}\left(\sum_1^\infty a_m^p\right)^{1/p}\left(\sum_1^\infty b_n^{p'}\right)^{1/p'}, \tag{12.6}$$

where

$$\begin{aligned} K_{m,n} &= \int_{m-1}^{m}\int_{n-1}^{n}\frac{dxdy}{x+y} \\ &= \int_{m-1}^{m}\{\log(x+n)-\log(x+n-1)\}\,dx \\ &= (m+n)\log(m+n)-2(m+n-1)\log(m+n-1)+ \\ &\qquad +(m+n-2)\log(m+n-2). \end{aligned} \tag{12.7}$$

Theorem 9 now follows from (12.3); and by a closely similar proof we obtain in the general, non-conjugate case.

THEOREM 10.

$$\sum_{m=1}^\infty \sum_{n=1}^\infty \frac{a_m b_n}{(m+n-1)^\lambda}\left\{1+\frac{2}{4!}\,\frac{\lambda(\lambda+1)}{(m+n-1)^2}+\frac{2}{6!}\,\frac{\lambda(\lambda+1)(\lambda+2)(\lambda+3)}{(m+n-1)^4}+\ldots\right\}$$
$$\leqslant \left(\pi \operatorname{cosec}\frac{\pi}{\lambda p'}\right)^\lambda \left(\sum_1^\infty a_m^p\right)^{1/p}\left(\sum_1^\infty b_n^q\right)^{1/q}.$$

13. Proofs that the constant $\pi \operatorname{cosec}(\pi/p)$, in the conjugate case of Hilbert's theorem, is best-possible are known [(1) 232]; but I have not discovered in the literature the rather simple proof given below.

Set $\qquad f(x) = x^{(\epsilon-1)/p}e^{-x}, \qquad g(y) = y^{(\epsilon-1)/p'}e^{-y};$

then, with $\epsilon > 0$,

$$\int_0^\infty f^p(x)\,dx = \frac{\Gamma(\epsilon)}{p^\epsilon}, \qquad \int_0^\infty g^{p'}(y)\,dy = \frac{\Gamma(\epsilon)}{p'^\epsilon}.$$

Hence

$$\pi \operatorname{cosec}\frac{\pi}{p}\left\{\int_0^\infty f^p(x)\,dx\right\}^{1/p}\left\{\int_0^\infty g^{p'}(y)\,dy\right\}^{1/p'} = \mathrm{B}\left(\frac{1}{p},\frac{1}{p'}\right)\left(\frac{1}{p}\right)^{\epsilon/p}\left(\frac{1}{p'}\right)^{\epsilon/p'}\Gamma(\epsilon). \tag{13.1}$$

Also

$$\begin{aligned}\int_0^\infty\int_0^\infty \frac{f(x)g(y)}{(x+y)}\,dxdy &= \int_0^\infty\int_0^\infty x^{(\epsilon-1)/p}y^{(\epsilon-1)/p'}e^{-(x+y)}\frac{dxdy}{(x+y)}\\ &= \int_0^\infty d\xi\int_0^1 \xi^{(\epsilon-1)/p}(1-\eta)^{(\epsilon-1)/p}\xi^{(\epsilon-1)/p'}\eta^{(\epsilon-1)/p'}e^{-\xi}\,d\eta,\\ &\qquad\qquad \text{where } x+y=\xi,\quad y=\xi\eta,\\ &= \int_0^\infty \xi^{\epsilon-1}e^{-\xi}\,d\xi\int_0^1 \eta^{\epsilon/p'+1/p-1}(1-\eta)^{\epsilon/p+1/p'-1}\,d\eta\\ &= \Gamma(\epsilon)\mathrm{B}\left(\frac{1}{p}+\frac{\epsilon}{p'},\frac{1}{p'}+\frac{\epsilon}{p}\right).\end{aligned} \tag{13.2}$$

That $\pi \operatorname{cosec}(\pi/p)$ is the best-possible constant now follows at once on comparing (13.1) and (13.2) and letting $\epsilon \to 0$.

REFERENCES

1. G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities* (Cambridge, 1934).
2. V. Levin, 'On the two-parameter extension and analogue of Hilbert's inequality', *J. of London Math. Soc.* 11 (1936), 119–24.
3. —— 'Two remarks on Hilbert's double series theorem', *J. of Indian Math. Soc.* 2 (1936–7), 111–15.
4. I. Schur, 'Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen', *J. für Math.* 140 (1911), 1–28.
5. S. B. Stečkin, 'On positive bilinear forms' (Russian), *Doklady Akad. Nauk SSSR* (N.S.), 65 (1949), 17–20.

ON SYMMETRIC RECURRENT TENSORS OF THE SECOND ORDER

By E. M. PATTERSON (*Sheffield*)

[Received 12 April 1950]

1. Introduction

In this paper, several general theorems are established for Riemannian spaces V_n which admit a second-order symmetric recurrent tensor T_{ij} other than a scalar multiple of the metric tensor g_{ij}. The tensor T_{ij} satisfies the equation

$$T_{ij,p} = \kappa_p T_{ij}$$

for some vector κ_i. Spaces of positive-definite metric admitting such a tensor are familiar, and it is known that they are reducible.† This paper is therefore concerned mainly with spaces of indefinite metric. These are not necessarily reducible; it is proved, however, that a space admitting a second-order symmetric recurrent tensor is reducible if the matrix (T^i_j) has at least two distinct latent roots at any point of the space.

It is shown that the existence of the tensor T_{ij} implies that there is a sequence of second-order symmetric recurrent tensors $T^{(\nu)}_{ij}$, corresponding to powers of the matrix (T^i_j). Each tensor $T^{(\nu)}_{ij}$, being recurrent, generates a parallel plane Π_ν. Sums and intersections of the parallel planes Π_ν and their conjugates give rise to further sets of parallel planes. It is found that the characters (dimensionality and nullity‡) of all these parallel planes are completely determined by the elementary divisors of the pencil $T_{ij} - \rho g_{ij}$ at any point of the space.

The Riemannian spaces considered in this paper are assumed to be covered by a single coordinate system. The components of the fundamental tensor g_{ij} possess derivatives up to any required order.

2. Recurrent tensors and parallel planes

A second-order covariant tensor T_{ij} $(i, j = 1,..., n)$ in V_n is *recurrent* if its covariant derivative $T_{ij,p}$ satisfies

$$T_{ij,p} = \kappa_p T_{ij} \tag{2.1}$$

for some vector field κ_p. In particular, κ_p may be identically zero, in

† See, for example, L. P. Eisenhart, *Trans. American Math. Soc.* 25 (1923), 297–306.

‡ By 'nullity' is meant the number of null vectors in a normal basis.

which case the tensor T_{ij} will be called *evanescent*. The fundamental tensor of V_n is evanescent. A V_n admitting a second-order symmetric recurrent tensor other than a scalar multiple of g_{ij} will be denoted by $\mathscr{T}_n$.

A set of p independent vector fields λ^i_α ($\alpha = 1,..., p$) defined over any V_n forms a basis for a *parallel p-plane*, i.e. a parallel field of p-dimensional vector-spaces, if

$$\lambda^i_{\alpha,j} = A^\beta_{\alpha j}\lambda^i_\beta \tag{2.2}$$

for some vectors $A^\beta_{\alpha j}$. Parallel planes have been discussed by A. G. Walker† and H. S. Ruse;‡ some of their results are quoted here.

Let Π be a parallel p-plane, and suppose that the vectors λ^i_α ($\alpha = 1,..., p$) form a normal basis (i.e. a set of mutually orthogonal vectors which are unit when non-null). If r of these vectors, say λ^i_ρ ($\rho = 1,..., r$), are null, and $p-r$ are non-null, then the vectors λ^i_ρ form a basis for a wholly null parallel r-plane. This is called the *null part* of Π, and is denoted by Π^*. Another parallel plane determined by Π is the *conjugate* $(n-p)$-plane Π', a covariant basis for which consists of the $n-p$ independent solutions of the equations

$$\mu_i\lambda^i_\alpha = 0 \quad (\alpha = 1,..., p).$$

If Π_1 and Π_2 are both parallel planes, then their sum $\Pi_1+\Pi_2$ and intersection $\Pi_1 \cap \Pi_2$ are also parallel planes. In particular $\Pi^* = \Pi \cap \Pi'$.

Suppose now that the V_n is a $\mathscr{T}_n$. If the symmetric recurrent tensor T_{ij} is of rank p (i.e. the matrix (T_{ij}) is of rank p), there are $n-p$ linearly independent solutions of the equations

$$T_{ij}\lambda^j = 0. \tag{2.3}$$

Because of (2.1), the solutions λ^i form a parallel $(n-p)$-plane, which is said to be *associated* with T_{ij}. The conjugate of the associated parallel plane is called the parallel plane *generated* by T_{ij}.

If the independent solutions of the equations (2.3) can be chosen so that they are all non-null and mutually orthogonal, the space admits a non-null parallel plane. It is then *reducible*; i.e. there is a coordinate system in terms of which the metric is§

$$ds^2 = g_{\alpha\beta}(x^\gamma)\,dx^\alpha dx^\beta + g_{\lambda\mu}(x^\nu)\,dx^\lambda dx^\mu$$

$$(\alpha, \beta, \gamma = 1,..., p;\ \lambda, \mu, \nu = p+1,..., n). \tag{2.4}$$

This is always the case when the metric is positive-definite and T_{ij} real.

† *Quart. J. of Math.* (Oxford) 20 (1949) 135–45.
‡ Ibid. 20 (1949) 218–34.
§ Eisenhart, loc. cit.

In a space $\mathscr{T}_n$, any tensor $T_{ij}^{(\nu)}$ of the sequence defined by

$$T_{ij}^{(\nu+1)} = T_{ik}^{(\nu)} T_j^k, \qquad T_{ij}^{(1)} = T_{ij} \quad (\nu = 1, 2, \ldots)$$

is a second-order symmetric recurrent tensor, satisfying

$$T_{ij,p}^{(\nu)} = \nu \kappa_p T_{ij}^{(\nu)}.$$

The process of forming the sequence is equivalent to taking powers of the matrix (T_j^i).

Since $T_{ij}^{(\nu)}$ is recurrent, it generates a parallel plane Π_ν. When the parallel plane generated by T_{ij} is non-null, each tensor $T_{ij}^{(\nu)}$ generates the same parallel plane. This is not true in the more general case of a partially null parallel plane, as will be shown below.

3. The elementary divisors of $T_{ij} - \rho g_{ij}$

In § 4 it will be shown that the dimensionality and nullity of the parallel planes Π_ν in a space $\mathscr{T}_n$ are determined by the number and nature of the elementary divisors to base ρ of $T_{ij} - \rho g_{ij}$. Before doing this, we show that the elementary divisors at any point are multiples of the elementary divisors at any other point at which T_{ij} does not vanish. This enables us to restrict attention to a single point of the space.

Let θ_α^i $(\alpha = 1, \ldots, n)$ be an ennuple of vectors parallel along a curve C in a space $\mathscr{T}_n$. The ennuple components of g_{ij} and T_{ij} with respect to θ_α^i are defined by

$$g_{\alpha\beta} = g_{ij} \theta_\alpha^i \theta_\beta^j, \qquad T_{\alpha\beta} = T_{ij} \theta_\alpha^i \theta_\beta^j. \tag{3.1}$$

Differentiating the second of these along C, we get

$$T_{\alpha\beta,k} \, dx^k/ds = T_{ij,k} (dx^k/ds) \theta_\alpha^i \theta_\beta^j,$$

and hence, since T_{ij} is recurrent,

$$dT_{\alpha\beta}/ds = \phi T_{\alpha\beta},$$

where $\phi = \kappa_i \, dx^i/ds$. Therefore $(T_{\alpha\beta})$ at any point Q of C is proportional to $(T_{\alpha\beta})$ at an arbitrary point P. In a similar manner it can be shown that $(g_{\alpha\beta})$ is constant along C. Hence the elementary divisors of $T_{ij} - \rho g_{ij}$ at any point are multiples of the elementary divisors at any other point at which T_{ij} does not vanish.

Since $$|T_{ij} - \rho g_{ij}| = |g_{kl}| |T_j^i - \rho \delta_j^i|,$$

the characteristic roots of T_{ij} are the same as the latent roots of (T_j^i), and the elementary divisors of $T_{ij} - \rho g_{ij}$ are the same as those of $T_j^i - \rho \delta_j^i$.

Walker† has shown that, if all the characteristic roots of T_{ij} are non-zero, then there exists a symmetric evanescent tensor of the form $\alpha(T_{ij}-\rho g_{ij})$, where α is a scalar and ρ is a characteristic root, so that the rank of the tensor is less than n. Hence any $\mathscr{T}_n$ admits a recurrent tensor S_{ij} such that (S_i^j) has at least one zero latent root.

Denote the matrix (T_i^j) at some point of the space by T. Then there exists a non-singular matrix ψ such that $\psi^{-1}T\psi$ is a matrix of the classical Jordan canonical form, i.e.

$$C=\begin{bmatrix} C_\sigma & . & . & . \\ . & C_\tau & . & . \\ . & . & C_\omega & . \\ . & . & . & U \end{bmatrix}, \qquad C_\sigma=\begin{bmatrix} . & 1 & . & . \\ . & . & 1 & . \\ . & . & . & 1 \\ . & . & . & . \end{bmatrix}$$

where a dot denotes a zero matrix or element. The square matrix C_σ is of order σ and corresponds to an elementary divisor ρ^σ. The matrix U corresponds to all the non-zero latent roots, so that $|U|\neq 0$. Since T has the same elementary divisors as C, it follows that the elementary divisors of $T_{ij}-\rho g_{ij}$ at any point of the space can be written down when the form of C is known.

4. The parallel planes Π_ν

THEOREM 1. *The dimensionality of the parallel plane Π_1' associated with T_{ij} is equal to the number of elementary divisors to base ρ of $T_{ij}-\rho g_{ij}$, and the nullity of Π_1' is equal to the number of non-simple elementary divisors.*

In order to prove this, we require the following two lemmas. The first is familiar, and follows at once from the form of the canonical matrix C. The second is probably well known, but a proof is outlined here because most standard works do not include one.

LEMMA 1. *If s is the number of elementary divisors to base ρ of $T-\rho I$, then T is of rank $n-s$.*

LEMMA 2. *In the vector space defined by the solutions of the equations $T_{ij}\lambda^j=0$, a unit vector of a normal basis corresponds to a simple elementary divisor of $T_{ij}-\rho g_{ij}$, and a null vector corresponds to a non-simple elementary divisor.*

Proof of Lemma 2. Let the symmetric matrices (g_{ij}) and (T_{ij}) at a particular point be denoted by G and H respectively. Then there exists a non-singular matrix P such that $J=P'GP$ and $K=P'HP$ form a

† Loc. cit., Theorem 7.1.

canonical pair† of matrices, with

$$J = \begin{pmatrix} J_{\sigma_1} & & \\ & J_{\sigma_2} & \\ & & \ddots \end{pmatrix} \qquad K = \begin{pmatrix} K_{\sigma_1} & & \\ & K_{\sigma_2} & \\ & & \ddots \end{pmatrix}$$

$$J_\sigma = \begin{pmatrix} . & 1 & . & . & \\ 1 & . & . & . & \\ . & . & . & 1 & \\ . & . & 1 & . & \ddots \\ & & & & \ddots \end{pmatrix} \qquad K_\sigma = \begin{pmatrix} . & . & . & & \\ . & . & 1 & & \\ . & 1 & . & \ddots & \\ & & \ddots & & \end{pmatrix}.$$

The sub-matrices J_{σ_α}, K_{σ_α} are each square matrices of order σ_α, and correspond to an elementary divisor ρ^{σ_α}. If σ_α is even, the last sub-matrices in J_{σ_α} and K_{σ_α} are respectively

$$\begin{pmatrix} . & 1 \\ 1 & . \end{pmatrix} \quad \text{and} \quad (\pm 1).$$

If σ_α is odd, they are

$$(\pm 1) \quad \text{and} \quad \begin{pmatrix} . & 1 \\ 1 & . \end{pmatrix}.$$

In matrix form, equations (2.3) become $H\lambda = 0$. If μ is a solution of $K\mu = 0$, then the corresponding solution of equations (2.3) is $\lambda = P\mu$. Since

$$\lambda_2' G\lambda_1 = \mu_2' P'GP\mu_1 = \mu_2' J\mu_1$$

for any two solutions μ_1 and μ_2 of $K\mu = 0$, it follows that a set of solutions $\mu_1, \mu_2, \ldots, \mu_\alpha, \ldots$ such that

$$\mu_\alpha' J\mu_\beta = 0 \quad (\alpha \neq \beta), \qquad \mu_\alpha' J\mu_\alpha = 0, \pm 1 \tag{4.1}$$

corresponds to a normal basis λ_α^i ($\alpha = 1, 2, \ldots$) for the vector space defined by the solutions of the equations $T_{ij}\lambda^j = 0$. The vectors μ_α satisfying $\mu_\alpha' J\mu_\alpha = 0$ correspond to null vectors of the normal basis, and those satisfying $\mu_\alpha' J\mu_\alpha = \pm 1$ correspond to unit vectors.

Define $\mu_1, \mu_2, \ldots$ by

$$\mu_1 = (1, 0, 0, \ldots),$$

$$\mu_\alpha = (0, 0, 0, \ldots, 1, 0, \ldots),$$

where the non-zero element occurs in the position $1+\sigma_1+\sigma_2+\ldots+\sigma_{\alpha-1}$. Then the vectors μ_α satisfy conditions (4.1), with $\mu_\alpha' J\mu_\alpha = 0$ if $\sigma_\alpha \neq 1$, and $\mu_\alpha' J\mu_\alpha = \pm 1$ if $\sigma_\alpha = 1$. Hence to each simple elementary divisor

† Bromwich, *Quadratic forms and their classification by means of invariant factors* (Cambridge Tract No. 3). It is possible to use these canonical matrices instead of the Jordan canonical form to prove all the theorems of the present paper.

to base ρ corresponds a unit vector of a normal basis for the solutions of the equations $T_{ij}\lambda^j = 0$, and to each non-simple elementary divisor corresponds a null vector. The lemma is therefore proved.

Returning now to the theorem, it follows from Lemma 1 that there are s independent solutions of the equations $T_{ij}\lambda^j = 0$, and therefore the dimensionality of Π_1' is s. From Lemma 2, the nullity of Π_1' is equal to the number of non-simple elementary divisors to base ρ. Theorem 1 is therefore proved.

THEOREM 2. *The parallel plane Π_ν' associated with $T_{ij}^{(\nu)}$ is of dimensionality $p_0 - p_\nu$ and nullity p_ν^*, where*

$$p_\sigma = e_{\sigma+1} + 2e_{\sigma+2} + 3e_{\sigma+3} + \ldots,$$
$$p_\sigma^* = e_{\sigma+1} + 2e_{\sigma+2} + \ldots + \sigma(e_{2\sigma} + e_{2\sigma+1} + \ldots),$$

and e_σ is the number of elementary divisors ρ^σ.

This theorem is proved by finding the elementary divisors to base ρ of $T^{(\nu)} - \rho I$ from those of $T - \rho I$, and the following lemma is required.

LEMMA 3. *Each elementary divisor ρ^σ of $T - \rho I$ gives rise to q elementary divisors ρ^{p+1} and $\nu - q$ elementary divisors ρ^p of $T^{(\nu)} - \rho I$, where $\sigma = p\nu + q$ and $0 \leqslant q < \nu$.*

Since $C^\nu = \psi^{-1} T^\nu \psi$, the elementary divisors of T^ν are the same as those of C^ν. In general, the latter is not a classical canonical matrix, but its elementary divisors to base ρ can be obtained by a suitable permutation of the columns and the same permutation of the rows. We have

$$C^\nu = \begin{bmatrix} C_\sigma^\nu & & & \\ & C_\tau^\nu & & \\ & & \ddots & \\ & & & C_\omega^\nu \\ & & & & U^\nu \end{bmatrix}, \qquad C_\sigma^\nu = \begin{bmatrix} \overbrace{\quad}^{\nu \text{ columns}} & & \overbrace{\quad}^{\sigma-\nu \text{ columns}} \\ & & 1 & & \\ & & & 1 & \\ & & & & \\ & & & & 1 \\ & & & & \\ & & & & \end{bmatrix}.$$

Then there exists a matrix χ such that

$$\chi^{-1} C_\sigma^\nu \chi = \begin{bmatrix} C_{p+1} & & & & \\ & C_{p+1} & & & \\ & & \ddots & & \\ & & & C_p & \\ & & & & C_p \end{bmatrix},$$

there being q sub-matrices C_{p+1} and $\nu-q$ sub-matrices C_p, where $\sigma = p\nu+q$. It follows that there are q elementary divisors ρ^{p+1} and $\nu-q$ elementary divisors ρ^p, and the lemma is proved.

Using Theorem 1, the dimensionality of Π'_ν is the number of elementary divisors to base ρ of $T^{(\nu)}_{ij}-\rho g_{ij}$, and the nullity of Π'_ν is the number of these which are non-simple. Let e_σ be the number of elementary divisors ρ^σ of $T_{ij}-\rho g_{ij}$. From Lemma 3, there are

$$e_1+2e_2+3e_3+\ldots+\nu(e_\nu+e_{\nu+1}+\ldots) = p_0-p_\nu$$

elementary divisors to base ρ of $T^{(\nu)}_{ij}-\rho g_{ij}$, and of these

$$e_{\nu+1}+2e_{\nu+2}+\ldots+\nu(e_{2\nu}+e_{2\nu+1}+\ldots) = p^*_\nu$$

are non-simple. Theorem 2 is therefore proved.

THEOREM 3. *If the matrix (T^i_j) has at least two distinct latent roots at any point, then the space $\mathscr{T}_n$ is reducible.*

Let ρ_1 and ρ_2 be two distinct latent roots of T. If T is non-singular, then there is a tensor $\alpha(T_{ij}-\rho_1 g_{ij})$ which has at least two distinct roots $\rho = 0$, $\rho = \alpha(\rho_2-\rho_1)$. If T is singular, take one of the roots ρ_1, ρ_2 to be zero. Let m be the greatest index of an elementary divisor to base ρ; then $e_\sigma = 0$ if $\sigma > m$. From Theorem 2, the parallel plane Π'_m associated with $T^{(m)}_{ij}$ is of dimensionality p_0 and nullity zero. If there is a zero latent root, $p_0 \neq 0$; and, if there is a non-zero latent root, $p_0 \neq n$. Hence the space $\mathscr{T}_n$ admits a non-null parallel p_0-plane which is neither the 0-plane nor the tangent n-plane; it is therefore reducible, being the product of a V_{p_0} and a V_{n-p_0}.

By a suitable choice of coordinates, in terms of which the metric takes the form given by (2.4), it can be shown that the space V_{p_0} admits a second-order symmetric recurrent tensor which has no non-zero characteristic roots. When all the characteristic roots are zero, $p_0 = n$; Theorem 2 can then be expressed in terms of the parallel plane generated by $T^{(\nu)}_{ij}$ as follows: *When all the characteristic roots of T_{ij} are zero, the tensor $T^{(\nu)}_{ij}$ generates a parallel plane of dimensionality p_ν and nullity p^*_ν.*

This parallel plane is wholly null if $p_\nu = p^*_\nu$. From the definitions of these numbers, this is true when $\nu \geqslant \frac{1}{2}m$.

5. Other parallel planes in a $\mathscr{T}_n$.

The sums and intersections of the various parallel planes Π_1, Π_2,... and their null parts are also parallel planes, all of which are contained in Π_1. In particular, the planes $\Sigma_\nu = \Pi^*_1 \cap \Pi^*_\nu$ are parallel.

THEOREM 4. *The null part of the parallel plane generated by T_{ij} contains a sequence of null parallel planes Σ_2, Σ_3,..., Σ_{m-1} such that $\Sigma_\nu \supset \Sigma_{\nu+1}$. The dimensionality of Σ_ν is $e_{\nu+1}+e_{\nu+2}+\ldots$ and $\Sigma_{m-1} = \Pi^*_{m-1}$.*

If $\lambda^i \in \Sigma_{\nu+1}$, then

$$T_{ij}\lambda^j = 0, \qquad T^{(\nu+1)}_{ij}\lambda^j = 0, \qquad g_{ij}\lambda^i\mu^j = 0$$

for every vector μ^i of $\Pi'_{\nu+1}$. Multiplying the first of these equations by $T^{(\nu-1)}_{kl}g^{ji}$, we get

$$T^{(\nu)}_{ij}\lambda^j = 0,$$

and, since any solution of $T^{(\nu)}_{ij}\xi^j = 0$ satisfies $T^{(\nu+1)}_{ij}\xi^j = 0$,

$$g_{ij}\lambda^i\xi^j = 0$$

for every vector ξ^i of Π'_ν. Hence $\lambda^i \in \Pi^*_\nu$, and therefore $\lambda^i \in \Sigma_\nu$, i.e. $\Sigma_{\nu+1}$ is contained in Σ_ν.

To prove the second part of the theorem, we use the canonical form C. Consider an elementary divisor ρ^σ ($\sigma > 1$) of $T_{ij}-\rho g_{ij}$. The solution λ of $C_\sigma\lambda = 0$ may be taken to be (1, 0, 0,...). This vector is also a solution of $C^\nu_\sigma\lambda = 0$, and by permuting the rows and columns of C^ν_σ as in the proof of Lemma 3, it can be seen that λ corresponds to a null vector of Π'_ν if and only if $\sigma > \nu$. Hence the number of null vectors in a basis for Π^*_1 which also occur in Π^*_ν is $e_\nu+e_{\nu+1}+\ldots$. In particular, for

$$\nu = m-1,$$

Σ_{m-1} is of e_m dimensions; since $\Sigma_{m-1} \subset \Pi^*_{m-1}$, and Π^*_{m-1} is of e_m dimensions, it follows that $\Sigma_{m-1} = \Pi^*_{m-1}$.

I should like to thank Professor H. S. Ruse and Professor A. G. Walker for their many helpful suggestions.

NOTE ON A THEOREM BY J. H. C. WHITEHEAD

By W. H. COCKCROFT (*Aberdeen*)

[Received 21 April 1950]

1. Introduction

LET K be a connected, C.W. complex,† with n-section K^n. In § 16 of (1), J. H. C. Whitehead proved that the second relative homotopy group $\pi_2(K^2, K^1)$ is a free crossed $(\pi_1(K^1), d)$-module‡, where

$$d\colon\ \pi_2(K^2, K^1) \to \pi_1(K^1)$$

is the boundary homomorphism into the fundamental group of K^1. The object of this note is to give a proof of this result which avoids the geometrical arguments, involving knot theory, which are used in (2).

2. Definitions

Let $K^2 = K^1 \cup \{e^2_\lambda\}$, where $\{e^2_\lambda\}$ is a set of 2-cells bounded by oriented circuits in K^1 and not meeting K^1 or each other in any point internal to any of them. Let

$$\rho_2 = \pi_2(K^2, K^1), \qquad \rho_1 = \pi_1(K^1)$$

with a common base point $x_0 \in K^1$, and, as above, let $d\colon \rho_2 \to \rho_1$ be the boundary homomorphism. Let $C_2 = H_2(\tilde{K}^2, \tilde{K}^1)$ be the second relative homology group, with integral coefficients, of the universal covering complex $\tilde{K}$ of K. Let $h_2\colon \rho_2 \to C_2$ be the homomorphism, defined in (2) 480, which is onto C_2 and whose kernel is the commutator subgroup of ρ_2.

3. THEOREM. *ρ_2 is a free crossed (ρ_1, d)-module.*

Let $a'_\lambda \in \rho_2$ be the element which is defined by a characteristic map for e^2_λ, joined to x_0 by a path in K^1. We define a free crossed (ρ_1, ϕ)-module h_π, with a basis $\{a_\lambda\}$, and with $\phi a_\lambda = da'_\lambda$. It follows from Lemma 2 of (2) § 2 that the correspondence $a_\lambda \to a'_\lambda$ determines an operator homomorphism, $\theta\colon h_\pi \to \rho_2$, such that $\phi = d\theta$. Obviously θ is onto ρ_2.

Let $c_\lambda \in C_2$ be the element which is represented by a characteristic map for $\tilde{e}^2_\lambda$, which is a cell in $\tilde{K}$ which covers e^2_λ. The set $\{c_\lambda\}$ is a natural basis for C_2, which is a free $\bar{\rho}_1$-module, where $\bar{\rho}_1 = \rho_1/d\rho_2 \approx \pi_1(K^2)$. The homomorphism $h_2\colon \rho_2 \to C_2$ carries a'_λ onto c_λ and is an operator

† For definition see (1) § 5.

‡ For definition see (2) 455, § 2.

homomorphism in the sense that $h_2(xa) = \bar{x}h_2(a)$, for all $a \in \rho_2$ and $x \in C_2$, where $\bar{x}$ is the co-set of $\bar{\rho}_1$ containing x. The correspondence $a_\lambda \to c_\lambda$ determines a similar operator homomorphism h: $h_\pi \to C_2$, onto C_2, with $h_2\theta = h$.

I now prove that θ is an isomorphism. Suppose $\theta a = 0$, then $d\theta a = 0$, therefore $\phi a = 0$. Also $h_2\theta a = 0$, therefore $ha = 0$. Hence, by Lemma 1 of (**2**), $a = 0$ and θ is an isomorphism.

REFERENCES

1. J. H. C. Whitehead, 'Combinatorial homotopy (I)', *Bull. American Math. Soc.* 55 (1949), 213–345.
2. ——, 'Combinatorial homotopy (II)', ibid. 453–96.

INFINITE MATRICES and SEQUENCE SPACES

Richard G. Cooke, D.Sc.

This book, by an acknowledged expert, gives a clear and detailed account of the theory of infinite matrices, developed with special reference to the summability of divergent sequences and series (from various points of view), with a discussion of reciprocals and bounds of infinite matrices, solution of certain linear matrix equations, Hilbert vector space, and other properties of infinite matrices employed in quantum mechanics.

42*s*.

★

THE ELEMENTS OF MATHEMATICAL ANALYSIS

Volumes I and II by

J. H. Michell, M.A., F.R.S.
and M. H. Belz, M.Sc., M.A.

In these two volumes Professors Michell and Belz have prepared a course of mathematical analysis which, while adapted to the particular needs of students of science and engineering, differs from other works of this type in that it is rigorous and on modern lines so far as mathematical principles are concerned. Little previous knowledge is assumed.

40*s*. each volume

MACMILLAN & CO. LTD.